Functional characterisation of MYT1L, a brain-specific transcriptional regulator

Kepa, Agnieszka Maria

Awarding institution:
King's College London

Functional characterisation of MYT1L, a brain-specific transcriptional regulator

Agnieszka Kępa

2014

A thesis submitted to King's College London for the degree of Doctor of Philosophy

MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry

ABSTRACT

Abnormalities in brain development and maladaptive plasticity are thought to underlay a range of neurodevelopmental and neurological disabilities and disorders, such as autism spectrum disorders, schizophrenia and learning disability. Though many transcription factors are known to control important diverse aspects of embryonic development, the molecular mechanisms by which these factors coordinate processes such as neuronal differentiation remain unclear. Thus, there is a need to identify genes and pathways involved in these processes. Myelin transcription factor 1- like (MYT1L) has a property to convert fibroblast to neurons and it is specifically expressed in the brain. Since the fundamental steps in neural development and *MYT1L* gene are highly conserved in vertebrates, we anticipated that interference with its function may result in obvious phenotypes. To understand the role of MYT1L in the vertebrate development, we first examined its expression patterns in the zebrafish and mouse. We reported *myt1l* first manifestation in the telencephalon region around 24h post-fertilisation and gradual increase of its mRNA levels during fish embryogenesis. We have also examined mRNA levels of this gene in mouse brain and found *Myt1l* most prominent role just before birth, when neural development is most active. To investigate the role of *MYT1L* during neural differentiation we used lentiviral-mediated gene silencing of this transcription factor in human neural stem cells. Subsequent microarray analyses revealed a list of potential targets of *MYT1L*, most of which were down-regulated upon *MYT1L* silencing, suggesting that the expression of those genes is mediated by *MYT1L*. The analyses of gene expression patterns during stem cells development have revealed that Myocyte Enhancer factor 2 (MEF2A)-specific binding site is present in most of genes co-expressed with *MYT1L*. We hypothesised, that *MEF2A* may regulate the expression of *MYT1L* and *MYT1L* co-expressed genes. By examining the consequences of *MEF2A* down-regulation in neuronal stem cells we discovered that *MYT1L* expression was negatively regulated upon *MEF2A* silencing. Differential gene expression analyses revealed that the majority of genes deregulated upon knock-down of *MEF2A* were also de-regulated by knock-down of *MYT1L*, suggesting that both genes operate through similar or possibly the same regulatory pathways. Therefore, the findings presented throughout this thesis contribute towards a better understanding of *MYT1L* role, function and its molecular mechanisms involved in the neural development. Further work is required to provide a greater understanding of *MYT1L* involvement in normal brain growth and maturation and the collective results might help to unravel molecular mechanisms underlying neurodevelopmental disorders and possibly help in identifying new therapeutic targets.

Table of Contents

TABLE OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr Sylvane Desrivières for the supervision and guidance offered throughout the duration of my study. I would also like to thank my second supervisor, Prof. Gunter Schumann, for giving me the opportunity to study for a PhD.

I am grateful to my colleagues in the Schumann lab, past and present, for their support and friendship, which made those long lab days much more enjoyable.

To my friends, especially Alanna, Hannah, Charlotte, Anna, Hannah and Grazynka thank you for your advice, support and for sharing those last few year with me.

I would like to express my sincere gratitude to Professor Francesca Happe for her ongoing help, encouragement and confidence in me. Completion of this work would not be possible without your help.

To Dad, thank you for your continuous support and believing in me and for your enormous help in many aspects of my life during those last few years. To Mum… thanks for watching over me. I hope I made you both proud

Statement of work

Chapter 1: Entirely my own work

Chapter 2: Zebrafish embryos were provided by Victoria Snowden (V.S.). Breeding, embryo maintenance and embryo preparation for experiments were performed by V.S. at the MRC Centre for Neurobiology, King's College London. Visualisation and image capture was done by Prof. Corinne Houart (C.H.). I have performed *in situ* hybridisation, RNA extractions, PCRs and qPCRs.

Chapter 3: My contribution included the generation of plasmids overexpressing MYT1L, tranfecting these plasmids into various cell lines as well as carrying out RNA analysis. shRNA MYT1L was provided by Lourdes Martinez Medina. Cell transduction, differentiation, RNA extractions and qPCR were performed by me. The RNA labelling, hybridisation, washing and scanning steps were outsourced and carried out by High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, Oxford. Microarray data analysis was performed by Dr Sylvane Desrivières. I performed GO-term enrichment analysis and IPA analysis.

Chapter 4: The zebrafish embryos were provided by V.S. I performed RNA extractions and qPCRs. I designed, produced and quantified shRNA MEF2A. I transduced and differentiated stem cells. I performed RNA extractions and qPCRs. The RNA labelling, hybridisation, washing and scanning steps were outsourced and carried out by High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, Oxford. Microarray data analysis was performed by Dr Sylvane Desrivières. I performed GO-term enrichment analysis and IPA analysis.

Chapter 5: Entirely my own work.

1. Introduction

Over the last decades, remarkable progress has been made in the field of neurobiology. In particular, advances in molecular biology have furthered our knowledge and understanding of how the brain develops and nerve cells communicate and how these internal cellular and molecular mechanisms relate to behaviour. In addition, the completion of the Human Genome Project and 'HapMap', have laid grounds for discovery of new genes for complex disorders.

Molecular genetics methods, such as genome-wide gene expression analysis, next generation sequencing, linkage and association studies will help us to understand genetic aspects of psychiatric diseases and integrate this findings to serve patients (Avramopoulos, 2010). Genetic and inherited components have been found for nearly all behavioural disorders, such as autism (Bailey et al., 1995; E. M. Morrow et al., 2008; Sutcliffe, 2008; Szatmari et al., 2007), schizophrenia (Bergen et al., 2012; Hamshere et al., 2013; Ripke et al., 2013), attention-deficit/hyperactivity disorder (ADHD) (Faraone et al., 2005; Silberg et al., 1996; Smalley et al., 2002; Smith et al., 2003; Stevenson, 1992) or intellectual disabilities (de Ligt et al., 2012; Gécz, Shoubridge, & Corbett, 2009; Plomin, Owen, & McGuffin, 1994; Rauch et al., 2012). Furthermore, genes and genetic mechanisms that may contribute to neurological disorders have been identified.

Besides this progress in the field of human genetics, the possibility to study model organisms, such as zebrafish, mice and primates, has allowed us to further our understanding of how different genes contribute to neurological abnormalities.

1.1. Normal brain development

Several major processes must be specifically coordinated during normal brain development. Any disruptions, due to either genetic or environmental factors, can result in cognitive deficits, leading to neurodevelopmental and psychiatric disorders, such as autism spectrum disorder (ASD) and schizophrenia (Ronan, Wu, & Crabtree, 2013).

Major stages of proper brain development are summarised below (see also figure 1.1).

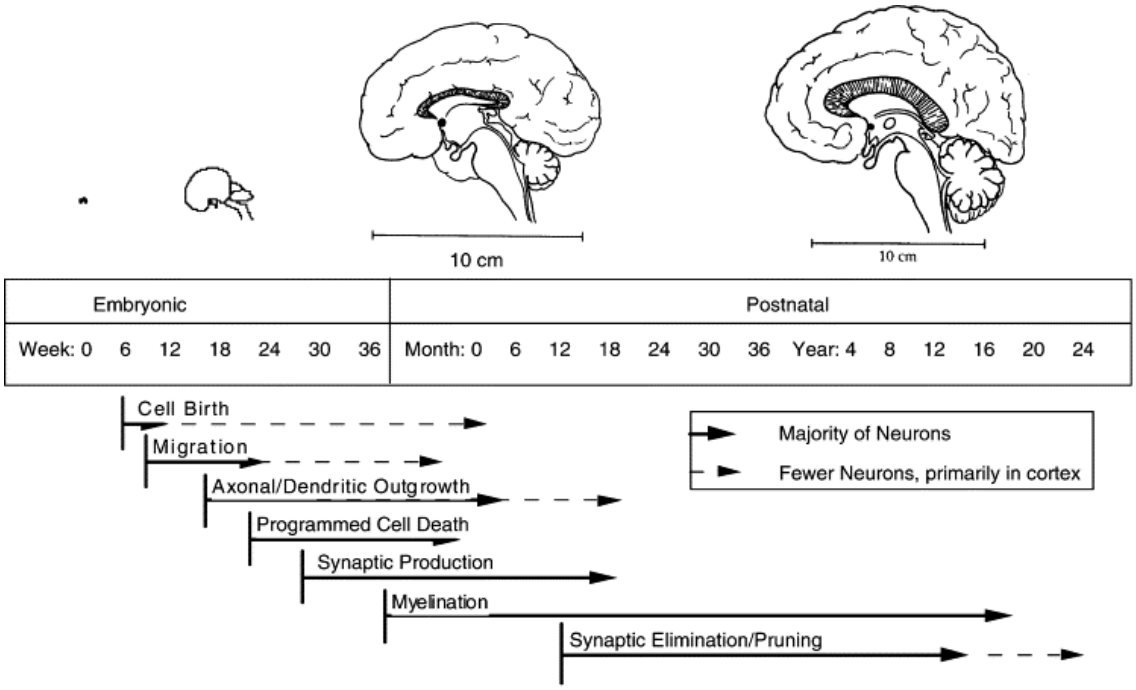


Figure 1.1. Major events in human brain development and times of their occurrence (picture from (Andersen, 2003).

1.1.1. Neurulation

Human brain development begins with the formation of the neural tube (neurulation) at gestational week three (Stiles & Jernigan, 2010). This consists of a primary and secondary stage, resulting in the formation of the central nervous system and spinal cord.

Primary neurulation spans from approximately the 18th gestation day to the end of the fourth week of gestation and it refers to the formation of the neural tube (Fotos, Olson, & Kanekar, 2011). This process is regulated by multiple genes and it involves complex morphogenetic events, orchestrated by cell proliferation, changes in cell adhesion and apoptosis (Back & Plawner, 2012). The nervous system originates from the dorsal point of the embryo as a neural plate which consists of tissue that differentiated in the middle of the ectoderm. At approximately day 18 of gestation the notochord and chordal mesoderm induce formation of the neural plate. Continuous induction directed by chordal mesoderm leads to dorsal closure of the neural plate and formation of the neural tube, which gives rise to the central nervous system (CNS). Disturbances during the events of primary neurulation are associated with a number of brain malfunctions. For example, fusion of the neural plate is regulated by the

paired-box 3 (*PAX3*) gene, in which any perturbations can lead to errors in neural tube closure (Sarnat, 2008).

Central to the fusion of the neural tube are two signalling molecules: bone morphogenetic proteins (BMPs) and Sonic hedgehog (Shh). They influence and induce the alar (a dorsal portion) and the basal (a ventral portion) plates that later form brain stem and spinal cord (Back & Plawner, 2012). Absence of sonic hedgehog will result in holoprosencephaly. The development and closure of the neural tube is usually completed by 28 days post-conception and failure of the posterior neural tube to close can result in spina bifida, while defects in the closure of the anterior neural tube could result in anencephaly (Blom, Shaw, den Heijer, & Finnell, 2006).

As the neural folds fuse, a population of specialised cells is formed from ectodermal cells on both sides of the neural tube. These cells form neural crest and they migrate shortly after the closure of the neural tube. They are precursor cells to melanocytes in the skin, facial connectivity tissues, sensory, sympathetic and parasympethtic ganglia, Schwann's cells and enteric neurons (Back & Plawner, 2012; B. K. Hall, 2008).

The process of secondary neurulation refers to the formation of the neural tube in the sacral and caudal regions from the caudal cell mass. At gestational days 41 to 51, the secondary neural tube and the central canal begins to regress together with the disappearance of the embryonic tail.

1.1.2. Prosencephalon development

The prosencephalon refers to the telencephalon and the diencephalon, future forebrain (Back & Plawner, 2012). The diencephalon later develops into the thalamus, hypothalamus and epithalamus, while the telencephalon gives rise to the cerebral hemispheres. Formation of the prosencephalon begins at the rostral part of the neural tube, peaking at 5-6 weeks after conception. The next stage, called prosencephalic cleavage, includes horizontal cleavage (formation of paired optic vesicles, olfactory bulbs and tracks), sagittal cleavage (to form the paired cerebral hemispheres, lateral ventricles and basal ganglia) and transverse cleavage (to separate the telencephalon from the diencephalon) (J. J. Volpe, 2000). During the next 2-3 months the midline prosencephalon develops to create the corpus callosum, the optic nerve and the hypothalamus. Disturbances in the formation of any of the aforementioned structures

can lead to a number of defects including holoprosencephaly which is a severe foetal brain malformation caused by a failure of cleavage of prosencephalon (Cohen Jr & Sulik, 1992)

1.1.3. Development of the cerebellum

The cerebellum is one of the most extensively studied parts of the brain. It integrates motor control with sensory perception, motor learning and cognition (C. C. Bell, 2002; Curtis C. Bell, Han, & Sawtell, 2008; Ito, 2002). The cerebellum is derived from the dorsal part of the anterior hindbrain and it develops over an extended period of time, starting at the early embryonic stage through to the first postnatal years. Its development begins after closure of the neural tube, and happens concurrently with prosencephalic development (ten Donkelaar & Lammens, 2009). Its growth occurs in four major steps:

Firstly, around 5^{th} and 6^{th} week of development, the two compartments of the cell proliferation are formed, the Purkinje cells and the deep cerebellar nuclei and later the Golgi, arise from the ventricular zone of the metencephalic alar plates. Next, towards the end of the embryonic period, granule cell precursors are formed from the upper rhombic lip. Granule precursor cells form the external granule layer, from which the cells migrate inwards through the Purkinje cells and form the internal granular layer of the mature cerebellum (ten Donkelaar & Lammens, 2009).

This complex and prolonged developmental process makes the cerebellum vulnerable to a broad spectrum of developmental disorders, such as spinocerebellar ataxia and Dandy-Walker malformation (Table 1.2) (Back & Plawner, 2012). Functional and anatomic abnormalities of the cerebellum have been linked to psychological disorders such as autism (Ito, 2008; Tsai et al., 2012).

1.1.4. Neuronal formation

The two main cell types that make up the nervous system are neuron and glia cells. The human brain contains billions of neurons, most of which are developed by mid-gestation (Bayer, Altman, Russo, & Zhang, 1993). Neurons, or nerve cells, are electrically excitable cells that process and transmit internal and external signals.

Major proliferative events occur between second and fourth month of gestation, with peak cell proliferation around 8 to 16 weeks (Figure 1.2).

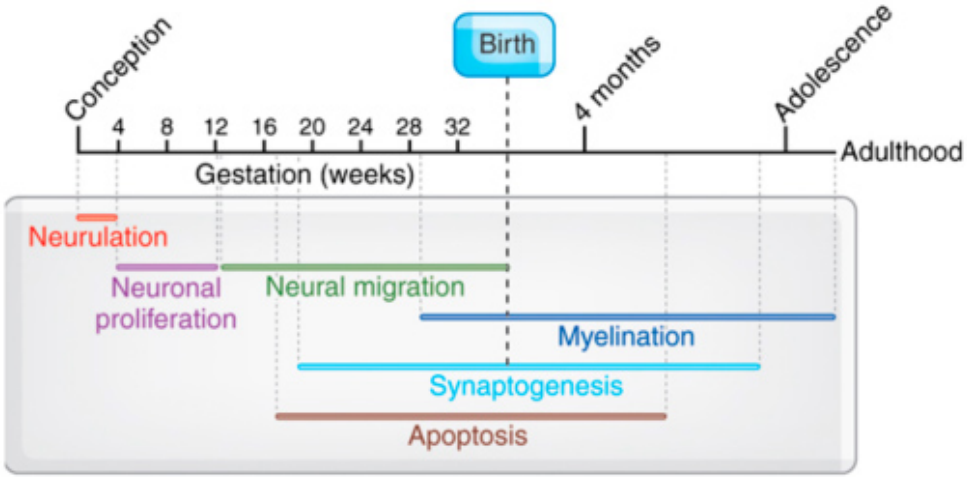


Figure 1.2. Timeline of the major events occurring during brain development. This figure represents brain development beginning with neurulation, and proceeding with neural proliferation and neuronal migration, synaptogenesis and pruning, which represent the creation and elimination of synapses during growth, myelination, and neuronal apoptosis (figure adapted from (Tau & Peterson, 2010)).

All neurons and glia are derived from ventricular and subventricular zones, present at every level of the developing nervous system (Joseph J. Volpe, 2008, pp. 55-118). With the closure of neural tube, at the end of gastrulation period, an area of neuroepithelial cells is formed. However, this group of neural stem cells is too small to sustain the whole creation of neurons, therefore the first step in production of neurons is to intensify and expand the populations of neural progenitor cells. Neural progenitor cells are mitotic cells, meaning they can divide and form new cells. The initial cycle of neural progenitor proliferation results from symmetrical expansion of the stem cell pool with each mitotic event creating two additional cells (Back & Plawner, 2012; Pasko Rakic, 1995). Over multiple rounds of cell division, the symmetrical expansion determines the total pool of stem cells from which cortical plate will form. Once the generation of stem cells pool is stable, the mode of cell division changes from symmetrical to asymmetrical (Figure 1.3). During this phase of clonal expansion, lasting from 5 months of gestation to 1 year after birth, two different cell types are produced, one neuronal progenitor cell and one neuron cell, that withdraws from the cell cycle (Caviness Jr & Takahashi, 1995).

Neurons are post-mitotic cells, meaning they cannot divide and produce new cells (Stiles & Jernigan, 2010; Wodarz & Huttner, 2003). The postmitotic neurons then begin migrating from the ventricular zone to the outer wall of the neural tube. The change to asymmetrical cell

division is gradual and as it progresses, proportionately larger number of postmitotic neurons, which are all derived from the same proliferative pool, and fewer stem cells, are produced.

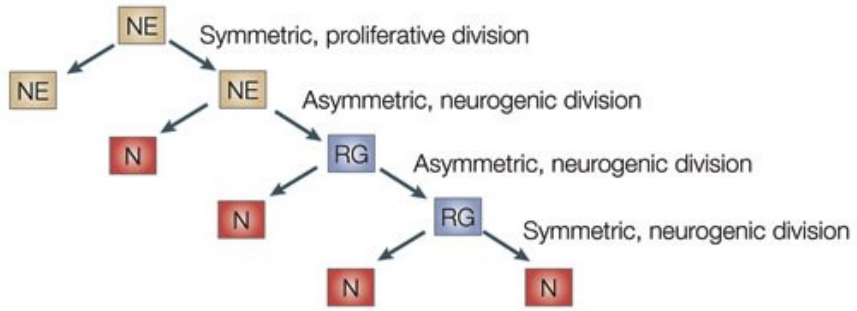


Figure 1.3. The lineage trees showing the relationship between neuroepithelial cells (NE), radial glial cells (RG) and neurons (N) (adapted from(Gotz & Huttner, 2005))

Disorders related to impaired neural proliferation would be expected to have a major influence on the development of the CNS and they include undersized brain (micrencephaly) and oversized brain (macrencephaly)(Joseph J. Volpe, 2008).

1.1.5. Neuronal migration

Neural migration refers to process by which millions of nerve cells move from their origin site in the ventricular and subventricular zones to their terminal sites across the CNS. Two distinct modes of migration have been identified so far: radial and tangential migration. Radial migration is the main mode of migration in the developing cerebral cortex. In tangential migration, neurons move parallel to the surface of the brain along axons or other neurons and often transgress regional boundaries. An example of this mode of migration is the movement of cortical interneurons from their origin in the ventral telencephalon to the developing cerebral cortex (Nadarajah & Parnavelas, 2002). In the early stages of brain development, neurons only travel very short distances, and the migration process is called somal translocation, which refers to a displacement of the cell body rather than migration of the whole cell (Nadarajah & Parnavelas, 2002). In somal translocation, neuron whose basal process (extension of the cell's body) is attached to the pial surface (the outer surface of the developing brain) and its cell body translocates as the process becomes shorter (Miyata, Kawaguchi, Okano, & Ogawa, 2001; Nadarajah, 2003). At the end of the somal translocation the nucleus of the cell has moved from the ventricular zone into the cortex.

With the developmental progress, the brain grows and the primary mode of neural migration changes. Rakic first proposed that young neurons, that have to travel greater distances, require 'radial glial guides' to support their migration (Figure 1.4) (P. Rakic, 1972; Sidman & Rakic, 1973). The radial glia processes extend from the ventricular zone into the pial surface of the brain and provide scaffolding along which neurons can travel. The migrating nerve cells attach themselves to the radial glia guide and move along the cellular scaffold into the developing cortical plate (Nadarajah & Parnavelas, 2002). Each glial scaffold can support the migration of many neurons.

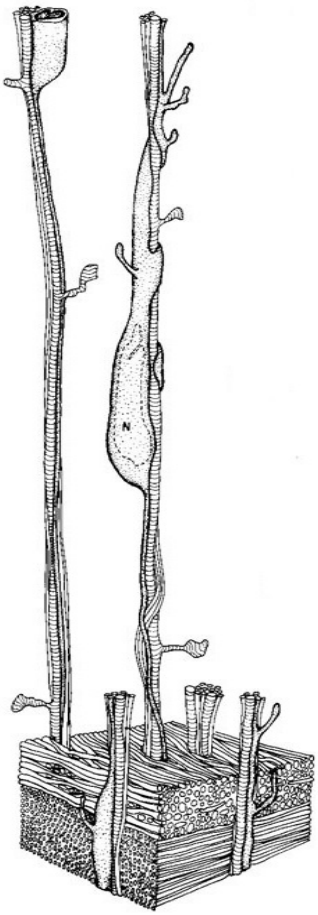


Figure 1.4. A three dimensional figure a migrating neuron (N) along the surface of a radial glial fiber (adapted from (P. Rakic, 1972)).

As neural migration advances, the preplate is split by the arrival of subsequent population of neuronal stem cells that will form the cortical plate. This results in the formation of the 6-layered cortical structure. The Cajal-Retzius cells of the marginal zone control the positioning of neurons into the correct layers of the cortex. Early-arriving neurons take deep position in

the cortex, and later arriving neurons take superficial positions, forming an 'inside-out' (from layer 6 to layer 1) sequence.

A wide spectrum of neural migration disorders can have a genetic background, such as schizencephaly, lissencephaly or pachygyria. Abnormalities of the *LIS1* (Lissencephaly 1), *DCX* (doublecortin), *ARX* (Aristaless-related homeobox), *TUBA1A* (Alpha-1A tubulin) and *RELN* (reelin) genes have been associated with these malformations (Guerrini & Parrini, 2010). Additionally, abnormal neuronal migration causes abnormal cortical function, which frequently results in cognitive and motor impairment and epilepsy (Guerrini, Dobyns, & Barkovich, 2008).

1.1.6. Neuronal differentiation and regionalisation.

Neural regionalisation events occur from around the fifth month of gestation to several years after birth. It is not known what specifies the regional identity of the different areas of the CNS (McKay, 1997). There are versatile expression patterns of cell surface signals and transcriptional regulators in the developing neuroepithelium before neuronal differentiation occurs (J. L. Rubenstein & Puelles, 1994). One hypothesis of how regional identity occurs could be that different stem cells for different brain regions. During neurogenesis, neurons are produced from almost all regions of the neuroepithelium, however it is unlikely that neurogenesis represents the default pathway for differentiation (Kintner, 2002). In general, neurons are produced before glia and specific type of each cell has also a specific time of 'birth'. Studies on *Drosophila* and grasshopper revealed that each stem-cell has its unique configuration based on its position within the neuroectoderm (Skeath, 1999). It has been suggested that stem cells fate can be at least partially due to intrinsic signalling (Shen et al., 2006), however how particular changes in progenitor characteristics occur remains largely unknown (Stiles & Jernigan, 2010). It is believed that the main genetic programming involves basic helix-loop-helix (bHLH) transcription factors, which are required for the differentiation of neuroepithelial cells into neurons, regardless of where and when they form. It is becoming increasingly plausible that most proliferating cells in the developing nervous system have been encoded with positional and patterning information that limits their developmental repertoire and proliferative potential and it is a stage-dependant change (S. Temple, 2001). Therefore, many progenitor cells are specified long time before their terminal differentiation.

Once they have reached their destination within a cortex, the neurons have to become a part of the information network. In order to communicate with each other, neurons need to

develop neuronal process (axons and dendrites). Neurons share a common organisation, controlled by their function, which is to receive, process and transmit information. Neurons exist in a variety of shapes and forms, but generally consist of four basic regions: dendrites, a cell body, an axon and axon terminals (Figure 1.5). Based on shape, neurons are classified into 3 major groups: unipolar, bipolar and multipolar (Figure 1.5). Unipolar neurons, are the most simple nerve cells, as they only have one primary process that can branch out. One branch serves as an axon, the other branches serve as dendrites. These type of neurons are present mainly in invertebrates; in vertebrates they are found in the autonomic nervous system. Bipolar neurons have a cell body that gives rise to two processes: one axon and a dendrite. Many sensory cells are bipolar cells, they are found in the retina and nose epithelium. Multipolar neurons are predominant in the vertebrate central nervous system. They have a single axon and typically number of dendrites that emerge from various points around the cell body. Multipolar cells vary greatly in their shape, especially in the axon length and the number of dendrites. Usually, the number of dendrites correlates with the number of synaptic connection the cell makes. Multipolar neurons predominate in the nervous system of vertebrates.

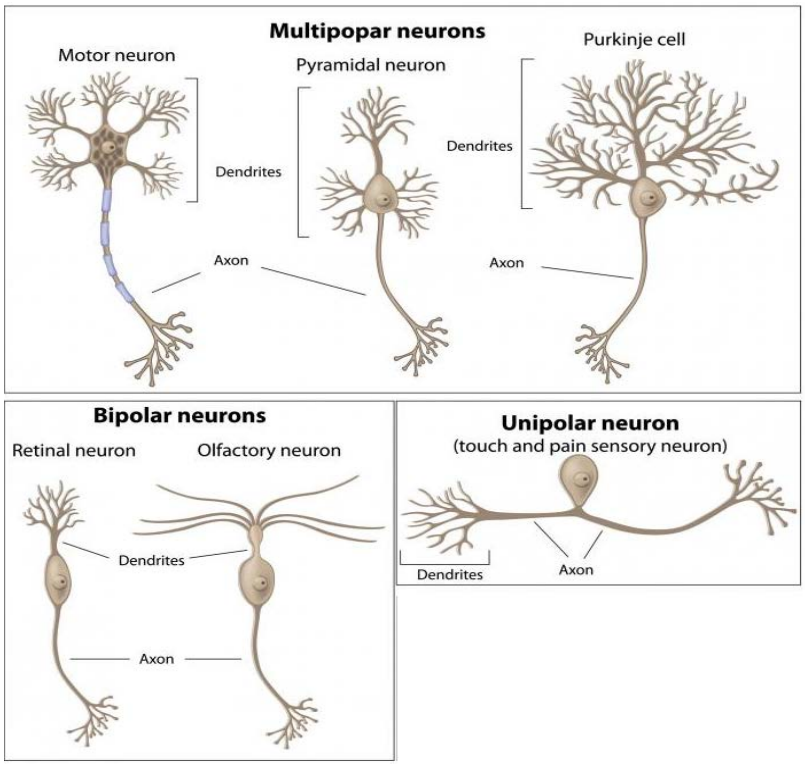


Figure 1.5. Different types of neurons based on their anatomical characterisation. (Adapted from Wisegeek.com).

Each part has distinct functions. Axons are the principal means of conveying the information over long distances and sending signals from the neuron to its terminals. Dendrites act as the receiver part of the neuron; their function is to process and transmit information to the cell body. This cell body, in which lies the nucleus, contains the genetic information that codes for production of cellular function elements. Each cell has multiple dendrites with a close proximity to the cell and a single axon that can extend for a long distance away from the cell. At the tip of each axon resides a growth cone, which is the site of axonal elongation and extension (Stiles & Jernigan, 2010). The growth cone acts as a guide that direct axon towards its target.

1.1.6.1. Synaptogenesis

Once the axon has reached its designated position, synapses are formed, which are the connections between the axon and the target cell. The function of the synapse is to transfer electric activity (information) from one cell to another. The nerve cell transmitting the information is called the presynaptic cell and the cell receiving a signal is the postsynaptic cell. For synapses to perform their task chemical neurotransmitters are released from the presynaptic neuron that bind and activate neurotransmitter-gated ion channels on the postsynaptic cell (Waites, Craig, & Garner, 2005). Activity-dependent and -independent process stimulates the initial steps of synapse differentiation and formation (Jiang et al., 1996). It has been shown that synapse number can be greatly regulated by non-neuronal signals, showing that glia may actively participate in synaptic plasticity (Ullian, Sapperstein, Christopherson, & Barres, 2001). Synapses are held together by cell adhesion molecules (CAMs) which stabilise the initial contact between axons and dendrites (Craig, Graf, & Linhoff, 2006). Research suggests that the actions of cell adhesion molecules is not only limited to providing adhesive support but they are also essential for subsequent multiple stages of synapse formation and maturation (Dalva, McClelland, & Kayser, 2007). A number of CAMs have been identified as mediating the assembly of the pre- and postsynaptic protein complexes. Neurexins and neuroligins were the first CAMs shown to be potent inducers for pre- and postsynaptic specialisation (Graf, Zhang, Jin, Linhoff, & Craig, 2004; Scheiffele, Fan, Choih, Fetter, & Serafini, 2000).

There are many types of synapses in the brain, each identified by the neurotransmitter they release. Examples include glutamate, acetylcholine, g-aminobutyric acid (GABA), glycine,

dopamine and serotonin. They also include neurotransmitters that are cell-specific, determining the phenotype and function of a neuron. The specificity of released neurotransmitters is mostly regulated by transcription factors (Goridis & Brunet, 1999). For example, two members of the forkhead/winged helix transcription factor family, Foxa1 and Foxa2, were found to be necessary to regulate the maintenance of dopaminergic neurons (Stott et al., 2013). The final stage of synaptogenesis is referred to as synapse elimination where inappropriate synapses are pruned via activity-dependent competition. The process is important for the formation of precise neural circuitry, which is necessary for proper brain functioning (Lichtman & Colman, 2000). During the early stages of development, synaptic contacts are generated in excess. During postnatal development the redundant synapses are eliminated while the proper ones are strengthened to construct specific neural connections (Kano & Hashimoto, 2009).

1.1.6.2. Apoptosis

Apoptosis, or programmed cell death is a genetically regulated and evolutionarily conserved process by which cells commit suicide. While most neurodevelopmental events involve differentiation and proliferation of the elements of central nervous system, cell death plays a critical role in the development of the nervous system.

Apoptosis occurs throughout the nervous system in neuron, glial, and neural progenitor cells. It is estimated that between 20 and 80% original cell population is eliminated as a result of apoptosis in the developing nervous system (Oppenheim, 1991). During neural development the role of apoptosis include optimisation of synaptic connections, removal of unnecessary neurons, and pattern formation (Burek & Oppenheim, 1999). Factors activating the cell death system seem to directly depend on the extent of neuron connections to a postsynaptic target which suggests that neurons are initially overproduced and then compete for target-derived neurotrophic factors, thus matching the size of the target cell population with the number of innervating neurons (Cowan, Fawcett, Oleary, & Stanfield, 1984). Failure of programmed cell death or over-activation of this process can have major detrimental effects for brain development and subsequent function. It has been suggested that inappropriate activation of apoptosis is responsible for some neuronal loss during stroke and trauma (Barinaga, 1998), Alzheimer's (Smale, Nichols, Brady, Finch, & Horton Jr, 1995), Huntington's (Hickey & Chesselet, 2003) and Parkinson's (Tatton, Chalmers-Redman, Brown, & Tatton, 2003) diseases.

1.1.7. Myelination

Though most of the production and differentiation of neurons occur prenatally, proliferation and migration of glial progenitor cells continues for extended period of time after birth (Stiles & Jernigan, 2010). Myelination refers to the process in which axons are wrapped with specialised myelin membrane that serves as an insulating layer and promotes a fast conduction of electrical impulses through the nerve cell. Myelin is synthesised by oligodendrocytes in the CNS, and it progresses most rapidly around and after birth (Umemori, Satot, Yagi, Aizawal, & Yamamoto, 1994). Initial proliferation and differentiation of oligodendrocytes is followed by expansion of the cell membrane to form the myelin sheath. Oligodendrocytes synthesise a number of trophic factors that play role in the axonal integrity and neuronal survival, therefore showing an oligodendrocyte influence on axonal diameter and neuronal size (McTigue & Tripathi, 2008).

1.2. Transcription factors and brain development

Important cell fate decisions during development of an organism are driven by gene expression changes, in which transcription factors play a pivotal role. Transcription factors (TFs) are proteins that control the expression of other genes. In general, they contain within their sequence variety of protein domains that bind to a specific consensus sequence of DNA and recruit additional cofactors to activate gene transcription (Nelson, 1995). Their activity is expected to play a crucial role in brain development and neural differentiation and functions, as these events depend on coordinated patterns of activation and inactivation of specific genes (He & Rosenfeld, 1991). Indeed, TFs play a key role in specifying cell identity during neuronal differentiation, and combination of transcription factors result in specific cell fates. These molecules perform their function alone or together with other proteins as part of complexes. To better understand different aspects of gene regulation, it is important to know the basic mechanisms behind this process.

Transcription can be divided into two parts: basal transcription and regulatory transcription. Regulatory elements can be located within the gene promoter, directly adjacent to the promoter region or at a distance from transcription start site (TSS) of the gene. A *cis*-regulatory region of the gene required for the start of the transcription is the designated promoter, which includes around 100bp including the transcription start site (Tamura, Konishi,

Makino, & Mikoshiba, 1996). A promoter is a target for basal transcription machineries. *Trans*-regulatory elements can lay millions of nucleotides away from the transcriptional start site and can interact with the *cis*- elements by directing transcription to specific cells or tissues.

The sequence to which a transcription factor binds with the highest affinity is known as its consensus binding site, and typically consists of 5-15 base pairs (Remenyi, Scholer, & Wilmanns, 2004). The core promoter directs initiation of transcription and contains several regulatory elements, such as basal transcription factors and RNA polymerase II which produce all protein coding and most non coding RNAs.

Because the primary function of transcriptional regulators is to govern the expression of target genes, the phenotypic abnormalities observed when TFs are misregulated are likely to be mediated by misexpression of target genes. Several neurodevelopmental disorders are caused by mutations of TFs (Hong, West, & Greenberg, 2005). For example, in the patients with schizophrenia, the expression of the transcription factor SRY-related HMG-box 10 (*SOX10*) was found to be decreased (Tkachev et al., 2003). Additionally, single nucleotide polymorphisms (SNPs) at the Neuronal PAS domain protein 3 *(NPAS3)* transcription factor gene locus was associated with increased risk of schizophrenia, bipolar disorder and major depression (J. Huang et al.; Sha et al.). Mutations within methyl CpG binding protein 2 *(MECP2)* gene are the cause of most cases of Rett syndrome (Shevell, 2009). disorders (Hong et al., 2005). Above examples illustrate the importance of transcription factors. However, how individual TFs contribute to brain development is largely unknown.

1.2.1. Myelin transcription factor 1-like (MYT1L)

One of the TFs that have been linked to brain development is the myelin transcription factor 1-like (MYT1L) gene product (J. G. Kim et al., 1997). MYT1L was first described recognised as a neural zinc finger transcription factor-1 (NZF-1) (Jiang et al., 1996). This gene is a member of the myelin transcription 1 gene family, which encodes neural specific, zinc-finger-containing DNA-binding proteins. In the human genome, there are three members of this family: myelin transcription factor 1 (*MYT1*), myelin transcription factor 1 like (*MYT1L*) and Myelin transcription factor 3 (*MYT3*), also known as Suppression of tumourigenicity 18 (*ST18* (J. G. Kim et al., 1997; S. J. Stevens et al., 2011; S. Wang et al., 2007). *MYT1L* is highly homologous to *MYT1*.

The protein MYT1 is a representative of the Cys-Cys-His-Cys (CCHC) zinc-finger protein family, which has been highly conserved during evolution between both species and family members. MYT3 is the third member of this family (Yee & Yu, 1998). Romm *et al.* suggested that MYT1/MYT1L binding can result in the recruitment of histone deacetylases (HDACs) to target gene promoters, which results in transcriptional repression (Romm, Nielsen, Kim, & Hudson, 2005). Jiang *et al.* first reported the presence of two separate zinc-finger DNA biding domains (Cys-Cys and His-Cys) in MYT1L, each of which could bind independently to similar DNA sequences (Jiang et al., 1996). MYT1 gene maps to human chromosome 20, while MYT1L gene is located in chromosome 2 in humans (2p25.3). The gene spans >542kb and has 25 exons. Exon 1 to 5 and a distal part of exon 25 are untranslated regions while the other 19 exons and proximal parts of exon 25 are coding regions. The *Myt1l* gene is localised to mouse chromosome 12. This localisation is consistent with that reported for NZF-1 (Jiang et al., 1996), indicating that the rat NZF-1 and the mouse Myt1l represent the same gene (J. G. Kim et al., 1997). Human MYT1L protein is similar to the Myt1l rodent protein, being 95% homologous to the mouse and 92% to the rat, showing its high conservation through evolution (S. J. C. Stevens et al., 2011).

Both MYT1L and MYT1 zinc finger proteins are found in neurons at early stages of differentiation (J. G. Kim et al., 1997). MYT1 was found to be highly expressed in cells of the subventricular zone (Armstrong, Kim, & Hudson, 1995), a germinal area from which neurons, astrocytes, and oligodendrocytes arise, suggesting that this transcription factor may play a role in the differentiation of neural progenitor cells (J. G. Kim et al., 1997). MYT1L, unlike MYT1, was not found in glial cells, but MYT1L protein was expressed in neurons after terminal mitosis (J. G. Kim et al., 1997). It was also suggested that transcripts encoding for MYT1L were present in differentiating neurons but not in neuroblasts. This suggested that MYT1L might have a role in the development of the nervous system. MYT1L is specifically expressed in the CNS (Berkovits-Cymet, Amann, & Berg, 2004; J. G. Kim et al., 1997). Additionally, the microarray data obtained online from the quantitative atlas BioGPS (available at http://biogps.org/#goto=welcome) has helped to provide evidence of the restricted expression of *MYT1L* in human brain (Figure 1.6) which is further evidence that it may play a role in the development of neurons in the brain.

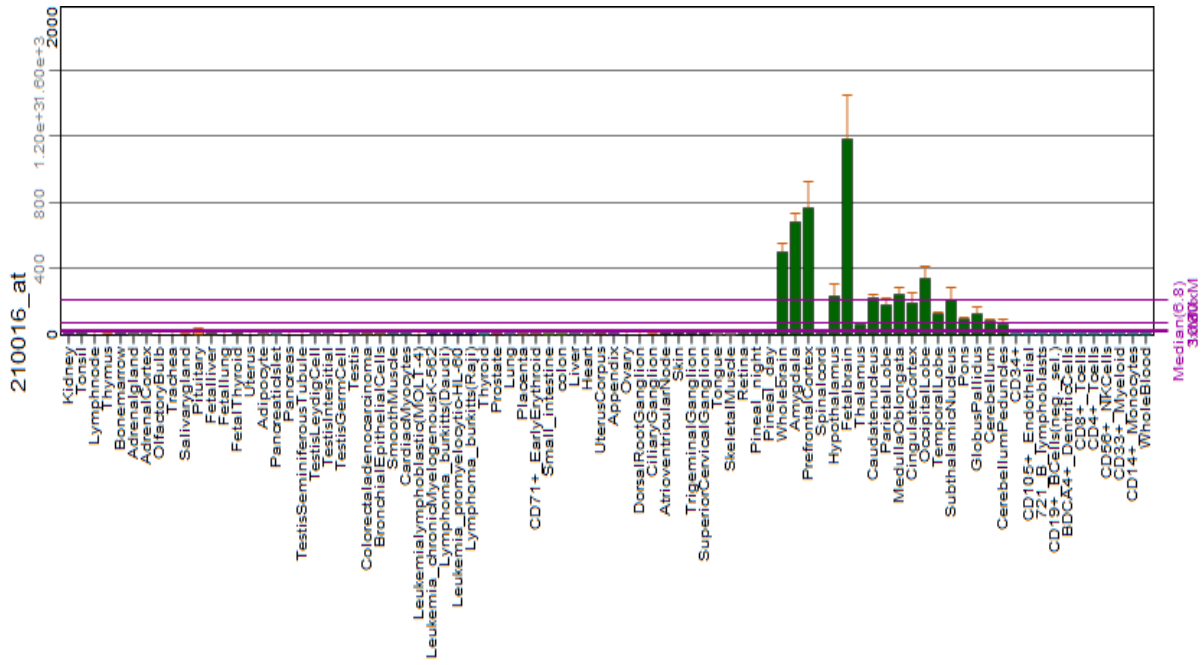


Figure 1.6. Expression pattern of *MYT1L* across different tissues and organs. The expression levels correspond to the results obtained from microarray experiments conducted with Affymetrix chips (Wu et al., 2009). To analyse human mRNA levels, the probe 210016_at on the Human Genome U133 array was used (Su et al., 2004); (source: http://biogps.org/#goto=welcome)

Jiang et al. (1996) first analysed the expression of rat Myt1l (called NZF-1). They found that NZF-1 is expressed in the rat's developing nervous system, primarily brain, spinal cord, sensory ganglia, retina, and nasal epithelia (Jiang et al., 1996).

The other family members such as Myt1 have also been characterised as brain specific (Armstrong et al., 1995). Myt1 is a zinc-finger DNA binding protein that was named for its ability to recognise the promoter region of the *proteolipid* (*PLP*) gene, the most abundantly transcribed CNS myelin gene (J. G. Kim & Hudson, 1992). During development of the oligodendrocytes, Myt1 is localised within nuclei of immature cells and then downregulated after terminal differentiation and accumulation of myelin proteins in mature oligodendrocytes (Armstrong et al., 1995). Myt1 continues to be expressed in germinal zones of the adult CNS and is upregulated in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas as well as in a dysembryoplastic neuroepithelial tumor (Armstrong et al., 1997) and following spinal cord traumatic injury (Wrathall, Li, & Hudson, 1998). *In vitro*, expression of a dominant negative form of Myt1 showed that Myt1 can regulate a critical transition in oligodendrocyte lineage cell development by modulating oligodendrocyte progenitor proliferation relative to

terminal differentiation and upregulation of myelin gene transcription (Nielsen, Berndt, Hudson, & Armstrong, 2004).

1.2.1.1. The importance of MYT1L during neuronal development

Brain specific expression of MYT1L and the fact that its expression is higher before birth suggests its crucial role in neurodevelopment (J. G. Kim et al., 1997). Recent studies showed that *Myt1l,* along with Achaete-scute homolog 1 (*Ascl1*) and POU class 3 homeobox 2 (*Brn2*), is capable of transforming fibroblasts into functional neurons (T. Vierbuchen et al., 2010). It was demonstrated that expression of *Ascl1* alone was sufficient to induce cells positive for markers of immature neurons (e.g., Tuj1, marker for neuronspecific β-tubulin). However, addition of *Brn2* and *Myt1l* was sufficient to induce cells with complex neuronal morphologies and functional synapses, thus indicating conversion to mature neuronal cells (T. Vierbuchen et al., 2010).

Next, the same three factors were shown to bgenerate functional neurons from human pluripotent stem cells. When those factors were combined with Neuronal differentiation 1 (*NEUROD1)* these factors could also convert fetal and postnatal human fibroblasts. These induced neuronal (iN) cells displayed typical neuronal morphologies and expressed multiple neuronal markers (Z. P. Pang et al., 2011). The *Ascl1/Brn2/Myt1l* complex alone appeared to induce immature neurons but was insufficient to generate functional neurons from human foetal fibroblasts. NEUROD1 aided the generation of the most mature neuronal cells. These studies were further corroborated when those iN cells were able to generate action potentials and many matured to receive synaptic contacts when co-cultured with primary mouse cortical neurons (Zhiping P. Pang et al., 2011).

It was also demonstrated that human fibroblasts could be reprogrammed to generate dopaminergic neurons. The factors Ascl1, Brn2, and Myt1l along with LIM homeobox transcription factor 1, alpha (*LMX1A*) and *FOXA2*, which are two transcription factors known to be involved in the developmental pathway for dopaminergic neurons, were necessary and sufficient for conversion. Induced cells were positive for tyrosine hydroxylase, an essential enzyme in the biosynthesis of dopamine and showed action potential activity when electrophysiology assays were performed (Pfisterer et al., 2011).

An alternative method was also proposed, in which gene silencing was used to directly trans-differentiate fibroblasts into neurons. MicroRNAs, miR-9/9*, and miR-124 expression in human fibroblasts induced their conversion into neurons, however addition of the previously known neurogenic transcription factors *NEUROD2*, *ASCL1*, and *MYT1L* enhanced the rate of conversion and the maturation of the converted neurons.

All of the aforementioned studies suggest a pivotal role of *MYT1L* in neuronal generation, differentiation and maturation.

1.2.1.2. MYT1L and disease

Considering the importance of *MYT1L* in neuronal development, it can be hypothesised that it may also play a role in the pathogenesis of certain neurodevelopmental disorders.

Wang et al. (2010) reported that that *MYT1L* can be a potential risk gene for major depressive disorder (MDD) in the Chinese Han population (T. Wang et al., 2010). They recruited 1139 patients and 1140 controls and used 8 SNPs as markers to investigate the role of *MYT1L* in MDD patients of Chinese origin. One SNP, rs3748989, located in exon 9, was found to be associated with MDD (T. Wang et al., 2010). This finding suggests that *MYT1L* may be a potential risk gene for MDD, although these results should be taken with caution as an obvious limitation to this study is the number of SNPs used; more markers should be utilised in subsequent studies for better mapping.

In a recent copy number variation (CNV) study, *MYT1L* was found to be disrupted in schizophrenia patients (Vrijenhoek et al., 2008). In this study, in the initial cohort of 54 schizophrenia patients, authors identified 13 rare CNVs. Each of the CNVs was detected in one patient only, illustrating the rare nature of these variants (Vrijenhoek et al., 2008). Among the genes disrupted by the rare CNVs was the *MYT1L* gene which has not been associated with schizophrenia before. Authors suggested that the (partial) duplications of *MYT1L* in patients 1 and 5 might affect *MYT1L* regulatory function in the CNS by either disruption or dosage effects (Vrijenhoek et al., 2008). Additionally, two patients with childhood onset of schizophrenia were reported to carry a microduplication disrupting the *PXDN* and *MYT1L* genes (Addington & Rapoport, 2009). As a replication, a meta-analysis of four published studies provided additional evidence for the association between microduplications in *MYT1L* and schizophrenia. The

overall rate of disruption in *MYT1L* found in this study was comparable to other CNVs that have been classified as high risk for schizophrenia (Yohan Lee et al., 2012)

The association of a *MYT1L* polymorphism and schizophrenia was also explored in the Han Chinese population. The results showed that two out of six SNPs studied, rs17039584 and rs10190125, had a significant association with schizophrenic patients in comparison to the controls. After dividing by gender, the latter SNP was significantly associated with SZ in female patients. These results must be taken with caution because they were only found in Chinese Han population and were not replicated in any other ethnic group.

A partial deletion within MYT1L has been associated with intellectual disability. In the case study of three siblings and three unrelated patients, a partial deletion of chromosome band 2p25.3 (2pter), ranging 0.37–3.13 Mb in size, was reported. All patients had intellectual disability. The only gene which overlapped and which was disturbed in all cases was *MYT1L*. It was therefore suggested that the cause of intellectual disability was due to *MYT1L* haploinsufficiency (S. J. Stevens et al., 2011).

Abberations in MYT1L have also been associated with autism. In this study, a *de novo* duplication in two male half-siblings with autism on chromosome 2p25.3, duplicating *PXDN* and partially duplicating *MYT1L* was reported. Their psychiatrically healthy mother was analysed and it was revealed that the transmission was due to germline mosaicism (Kacie J. Meyer, Michael S. Axelsen, Val C. Sheffield, Shivanand R. Patil, & Thomas H. Wassink, 2012).

An important role for *MYT1L* in normal brain development is supported by the aforementioned studies. As described, first links between *MYT1L* and human neurodevelopmental disorders have been provided. However, the precise mechanisms through which MYT1L acts and its downstream target genes remain unknown warranting further investigation into its function.

1.3. Aims and objectives

The overall aim of the present project was to elucidate and characterise genetic functions of the transcription factor - MYT1L. Recent publications have identified a network consisting of *Ascl1, Brn2* and *Myt1l* to be sufficient to convert human and mouse fibroblast into function neurons (Pfisterer et al., 2011; T. Vierbuchen et al., 2010). The literature available suggests it has role during neuronal cell differentiation, maturation and possible synapse development.

Using expression quantitative trait loci (eQTL) mapping data, which offer a unique opportunity to evaluate the impact of genetic variation on gene expression, the supervisor's laboratory have analysed gene expression profiles associated with and encoding for MYT1L (data unpublished). It has been found that single nucleotide polymorphisms (SNPs) in *MYT1L* are associated with expression of over 2000 transcripts, suggesting that these genes might be downstream targets for MYT1L. eQTL analysis provides a link between that particular SNP and those affected genes by measuring their expression levels. Pathway analysis suggested that *MYT1L* regulates a large number of genes involved in neural differentiation and neurite development. These findings imply that MYT1L may play a central role in brain development

The aim of this study was sub-divided into following objectives:

1. To elucidate the expression patterns of *Myt1l* during vertebrae embryonic development, with zebrafish and mouse used as model organisms.

2. To understand the role of *MYT1L* in neuronal cell differentiation, the consequences of its overexpression and down-regulation were examined *in vitro*, using human neural stem/progenitor cell lines as a model for neural differentiation.

3. Analysis of *MYT1L* co-expressed genes provided us with a potential candidate gene that can act as a regulator of *MYT1L* and its network – *MEF2A*. Therefore, the aim of the last part of this thesis was to further elucidate the link between *MYT1L* and *MEF2A*. To address this question, lentiviral mediated *MEF2A* silencing in neural stem cells was implemented, followed by gene expression analysis.

2. Role of *myt1l* in vertebrate brain development

2.1. Introduction

Animal models play a central role in the scientific investigation of normal and abnormal behaviours, and for the study of the (patho) physiological mechanisms underlying these behaviours (Fisch, 2007; Matthews, Christmas, Swan, & Sorrell, 2005; Petters & Sommer, 2000; Phillips et al., 2002). Understanding the functions of many genes that are necessary for development is one of the central goals in biology. Researchers have used genetic approaches in model organisms to dissect how neurodevelopmental processes work. A number of studies have shown that developmentally important genes are conserved throughout species (Bier & McGinnis, 2004; McCarroll et al., 2004; Santini, Boore, & Meyer, 2003), suggesting that developmental processes in humans will be controlled by a similar sets of genes in vertebrates. However, there is a critical gap in our knowledge of how these genes act and what pathways and processes they regulate in the physiological setting, which can be filled through experimental and hypothesis-driven approaches in model organisms (Mouse Genome Sequencing et al., 2002).

Several animal models can be used to study and analyse the function of a given gene. Here, I present data from two animal models that can be used to detangle genetic involvement during development. I used zebrafish and mouse as organisms of choice, due to features that are described below. Little is known about molecular and biological functions of MYT1L in human, thus choice of distant but related vertebrates to examine genetic properties of this transcription factor is plausible.

2.1.1. Zebrafish as a model organism

Zebrafish (*Danio rerio*) has been introduced as a model for development and genetic research by George Streisinger in 1981 (Streisinger, Walker, Dower, Knauber, & Singer, 1981; Walker & Streisinger, 1983). This animal has been widely studied as a model organism for vertebrate development because it appears to combine the best features of all the other models (Twyman, 2002). Though a vertebrate, it shares many strengths with invertebrate models, such as small size, easy maintenance and large number of progeny produced in a short time, providing physiological and morphological complexity of chordate animal systems (Goldsmith,

2004). Zebrafish embryos develop externally from the mother's body, therefore they can be viewed and manipulated at all stages. These are also transparent and show rapid development, taking only 5 days from fertilisation to fully swimming larva (Charles B. Kimmel, Ballard, Kimmel, Ullmann, & Schilling, 1995). Zebrafish are smaller than mice and they produce more offspring in a shorter time (Twyman, 2002). Considering all of the aforementioned aspects, the zebrafish is an ideal vertebrate model to study embryonic development (Figure 2.1). The embryonic development is described in six broad periods of embryogenesis and each period is subdivided into stages, which are named according to their morphological features: the zygote (1-cell), cleavage (2-cells to 64-cells), blastula (128-cells to 30% epiboly), gastrula (50% epiboly to bud), segmentation (1-4 somites to 26+ somites) and pharyngula (prim-5 to long-pec) periods. During segmentation, somitogenesis occurs, which is the formation of somites from the presomitic mesoderm. The 16 somite stage marks the onset of ventral bending of the neural axis at the level of the cephalic flexure, a major morphogenetic process that generates the displacement of forebrain territories at distinct axial levels. The earliest dopaminergic (DA) neurons in zebrafish are detected at ≈24 h postfertilisation (hpf) in the basal forebrain (Guo et al., 1999).

Zebrafish has been a subject to large-scale genetic screening, which identified thousands different mutations affecting development (Blader & Strähle, 2000). These huge screenings provided a further affirmation that zebrafish is a powerful genetic system and they allowed studying gene functions and complex developmental processes regulating embryogenesis in vertebrates (Driever et al., 1996; Haffter et al., 1996).

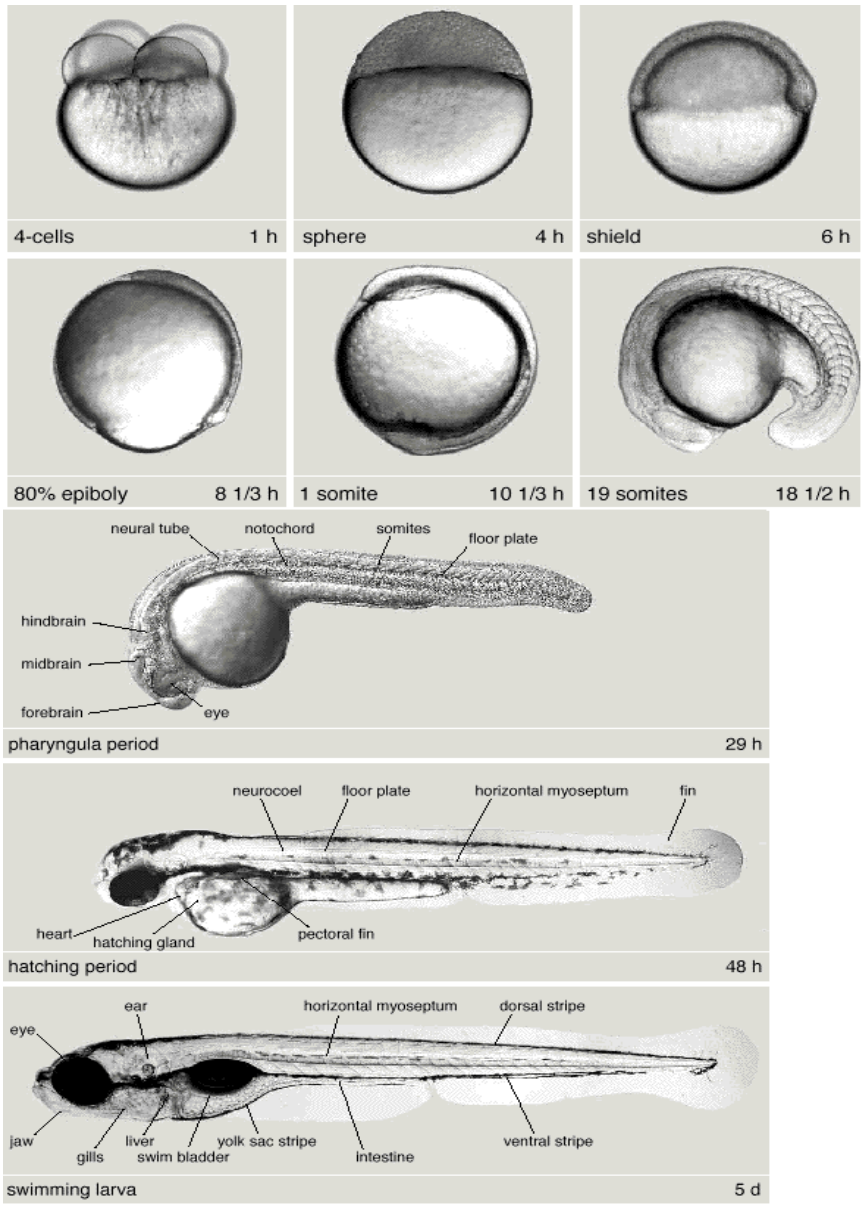


Figure 2.1. Embryonic development of the zebrafish. Pictures represent different stages of the developing embryo. Characteristic structures visible during embryogenesis are indicated. Stages are according to (Charles B. Kimmel et al., 1995). Picture adapted from (Haffter et al., 1996; Nüsslein-Volhard, 1995).

With the development of more and more cutting edge technology to study molecular processes, researchers can benefit from a wide variety of tools and information on how to use zebrafish as a model organism. It has been established as an important vertebrate model to study human disease including heart disease (Chico, Ingham, & Crossman, 2008), cancer (Amatruda & Patton, 2008), motor neuron disease (Beattie, Carrel, & McWhorter, 2007) and Alzheimer's disease (Newman, Musgrave, & Lardelli, 2007). Furthermore, it is being used as a tool for high-throughput toxicology and drug discovery screens (Kari, Rodeck, & Dicker, 2007).

The zebrafish system has also been presented as a potential viable model for neurodevelopmental diseases, such as autism and autistic disorders (Tropepe & Sive, 2003) and schizophrenia (Morris, 2009). One avenue that can effectively model some of the abnormalities observed in individuals with neurodevelopmental disorders, is based on a candidate gene approach.

There are established molecular genetic techniques to study the function of susceptibility genes in zebrafish.

One of the methods to investigate loss of gene functions involves morpholino oligonucleotides (MOs) (Bill, Petzold, Clark, Schimmenti, & Ekker, 2009). MOs are chemically modified oligonucleotides that bind to the targeted transcript and interfere with RNA translation (Nasevicius & Ekker, 2000). To interfere with genetic function, MOs are injected in the early stages of embryonic development (for example 1/128 blastula stage) and consequences on fish development can be analysed by *in situ* hybridisations, immunocytochemistry or live imagining. As a consequence of embryo transparency, individual cells can be clearly visualised *in vivo* using microscopic techniques across a broad range of developmental stages. Whole mount *in situ* hybridisation (ISH) is a commonly used technique to determine gene expression patterns during early development. During the *in situ* hybridisation procedure, an antisense mRNA probe is designed to recognise and bind the endogenous transcript, which is later detected by the color-based or fluorescence-based assay.

Sequencing of the zebrafish genome is close to completion with the latest release, Zv9 whole genome assembly (http://www.ensembl.org/Danio_rerio/Info/Index). Therefore, making it the largest gene set of any vertebrate sequenced so far (Howe et al., 2013). Zebrafish possess over 26000 protein-coding genes (Collins, White, Searle, & Stemple, 2012), more than any previously sequenced vertebrate, and they have a higher number of species-specific genes in their genome than do human, mouse or chicken (Howe et al., 2013). Although, there is a large homology between zebrafish and human genes, there are numbers of genes that underwent duplication during evolution. As a result, zebrafish often have two co-orthologs in contrast to a single copy gene in humans and other mammals (Taylor, Braasch, Frickey, Meyer, & Van de Peer, 2003; Woods et al., 2005). These duplicates may have different expression patterns and/or novel or unrelated functions (Morris, 2009). As a consequence, the function of individual genes might be less complex in fish than in tetrapods, therefore easier to study in the former. myt1l is highly conserved throughout vertebrate development (S. Wang et al., 2007) with 77% homology between human and zebrafish proteins. This suggests that it

controls a biological process that has been preserved during vertebrate evolution. In mammals there is only one *MYT1L*, while in zebrafish there are two paralogs of this gene – *myt1la* and *myt1lb*. According to zebrafish genome browser (zfin.org), myt1la (Gene ID: 559505) is the only form that has been characterised as protein coding and it is an ortholog of human, mouse and rat Myt1l. It maps to chromosome 20, where myt1lb (GI:23188298) has not got any genetic coordinates or any known biological function. It has only been identified by statistical methods as a part of large EST mapping project (Lo et al., 2003). Due to lack of information and possible nonfunctionalisation of myt1lb, this thesis concentrated only on characterisation of myt1la as an ortholog of mammalian Myt1l.

With its many advantages as a vertebrate model to study development, combined with genetic screening and manipulations, zebrafish is becoming a very interesting and efficient model that can be used to search for genes involved in neurodevelopment (Fishman, 1999; Vascotto, Beckham, & Kelly, 1997).

2.1.2. Mouse as a model organism

Over the past years, the mouse has become one of the most used mammalian model system for genetic research. Scientists from a wide range of biomedical fields have been using mouse because of its close genetic and physiological similarities to humans, as well as the ease with which its genome can be manipulated and analysed. With both human and mouse genomes sequenced, it has been established that 99% of encoded genes are shared between the two (Mouse Genome Sequencing et al., 2002). Practically, mice are a cost-effective and efficient tool for development of new drug therapies.

Mice are small, have a short generation time and an accelerated lifespan (one mouse year equals about 30 human years), keeping the costs, space, and time required to perform research manageable (Kile & Hilton, 2005). The sequencing of the mouse genome and the characterisation of a range of strain specific genetic markers allow mutations and genetic alterations to be readily mapped (Mouse Genome Sequencing et al., 2002).

There is an obvious difficulty with obtaining early human embryonic and foetal material, making mouse models very useful for the study of mammalian development and also for the collection of many mammalian genes in the form of cDNAs. Mouse development *in utero* lasts

for 20 days, whereas human development lasts for 280 days, and the corresponding stages of development are not evenly distributed throughout the gestation period (Ko, 2001). For example, embryonic day 10.5 (E10.5) is the period of development characterised by closure of the caudal neuropore (Kaufman, 1992). Stage E13.5-E14.5 is associated with neurogenesis and the beginning of cell migration to the primary neuropallial cortex. At E17.5 the brain displays an increased degree of differentiation (Schurov, Handford, Brandon, & Whiting, 2004). Altogether, their detailed characterisation, convenience of breeding and relatively close evolutionary relation to human has ascertained that mice are one of the best resources to model human development and disease (Hacking, 2008). For the purpose of our study, I interrogated the Allen Mouse brain atlas for expression details on Myt1l (Figure 2.2). These data shows that *Myt1l* is not expressed at E11.5 and its expression starts at E13.5 and continues postnatally until day P56.

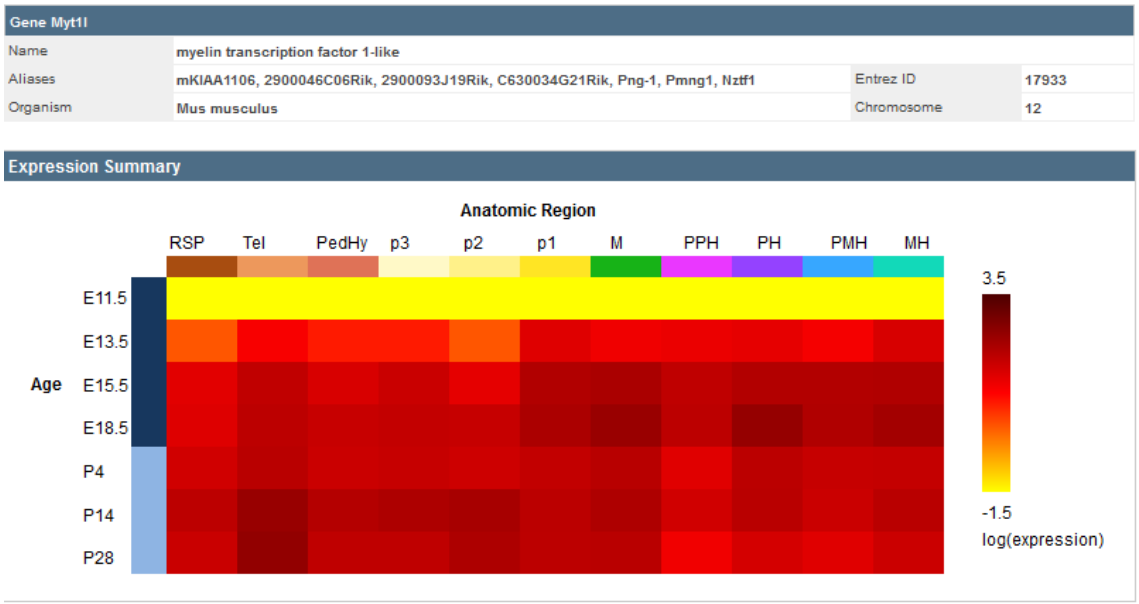


Figure 2.2. Summary of expression levels of *Myt1l* during mouse brain development. Dark colours represent higher expression and light colours represent no or little expression. Embryonic stage E11.5 shows no expression with *Myt1l* detection being recorded beginning of stage E13.5 in all tested brain regions. RSP - Retrosplenial area. Tel – telecephalon, PedHy- peduncular (caudal) hypothalamus, p1-3 – prosomere 1-3, M- midbrain, PPH- prepontine hindbrain, PH- pontine hindbrain (pons proper), PMH- pontomedullary (retropontine) hindbrain, MH- medullary hindbrain (medulla) (from Allen Developing Mouse Brain Atlas. Available from: http://developingmouse.brain-map.org)

2.2. Aims and hypothesis

Myt1l is specifically expressed in the human brain, as mentioned already in Chapter 1. Since the fundamental steps in neural and brain development are highly conserved in vertebrates, we anticipate that Myt1l function and expression pattern will also be conserved. I propose to use zebrafish to examine *myt1l* expression patterns during embryonic development. The principal aim of this study was to map the distribution of *myt1l* mRNA and during embryonic and early larval stages of zebrafish development. Specific techniques developed for zebrafish, such as *in situ* hybridasation are proposed to investigate the role of *myt1l* during embryonic development. Using qPCR, changes of *myt1l* expression during fish maturation were assessed.

Another aim was to validate zebrafish findings in a higher vertebrate group – mice, and to investigate Myt1l mRNA expression profiles, by qPCR during mouse brain development. To achieve this, several pre-natal (E10, E14, E18) and post-natal (P1week, P1month, P6months) stages of mouse brain development were analysed.

2.3. Materials and methods.

All zebrafish studies were performed with approval from the UK Home Office under a HO project license to Professor Corinne Houart (C.H.). Breeding, embryo maintenance and embryo preparation for experiments (such as fixation and dehydration) were performed by Victoria Snowden (V.S.) at the MRC Centre for Neurobiology, King's College London. All embryos were provided by V.S. Visualisation and image capture was done by C.H. I have performed *in situ* hybridisation, RNA extractions, PCRs and qPCRs.

Breeding fish were maintained at 28.5°C on a 14-hour light/10-hour dark cycle. Embryos were staged according to the protocol described previously by Kimmel *et al.* (Charles B. Kimmel et al., 1995). Collected embryos were cultured in fish water containing 0.003% 1-phenyl-2-thiouera to prevent pigmentation. Embryos were allowed to develop in regular fish water until the end of gastrulation. For embryos older than 24 h, in order to prevent pigmentation, fish water was replaced, at the end of gastrulation by a solution of 0.0045% 1-phenyl-2-thiourea and 0.01% methylene blue to prevent fungal growth.

2.3.1. Zebrafish whole mount in situ hybridisation (ISH)

To investigate expression patterns of *myt1l* during embryonic development we used *in situ* hybridisation. Whole mount *in situ* hybridisation (ISH) is commonly utilised to determine gene expression patterns during early development by detecting specific nucleic acid sequences with RNA probes. ISH consists of few steps, detailed below and previously described by [Thisse and Thisse. (C. Thisse & Thisse, 2008). Briefly, a full length clone of zebrafish MYT1L inserted in pCR-4-TOPO was used as a template for the synthesis of an antisense RNA probe, which was labelled with digoxigenin-linked (DIG) nucleotides. Following this, embryos collected at various stages of development, were fixed and permeabilised before being soaked in the digoxigenin-labelled probe. Hybrids (embryos with DIG-labelled RNA probes) were detected by immunohistochemistry using an anti-DIG antibody.

2.3.1.1. Preparation of antisense DIG labelled RNA Probes

A zebrafish cDNA *myt1l* clone was purchased through Open Biosystems (clone ID: 9039342, accession: BC171626). The clone was obtained in pCR 4-TOPO (Figure 2.3) vector that already contains a T7/T3 priming site that is necessary for *in situ* hybridisation. The clone's length is 3872 bp and it represents a full length, open reading frame (ORF) clone. Manufacturer provided the clone in a bacterial stock. To prepare the DNA plasmid the sterile tip was dipped into the vial containing stock, and the bacteria were allowed to proliferate in a pre-culture of 5 ml of LB Broth media (Invitrogen, UK) containing 100 µg/ml ampicillin (prepared as described in 3.3.1.4) at 37^{o}C overnight and used for large-scale plasmid production. Plasmid DNA for was isolated by the Endofree Plasmid Maxiprep Kit (Qiagen, UK) (as described in 3.3.1.8). DNA concentration was determined using a NanoDrop ND-1000 spectrophotometer and plasmids were stored at -20^{o} C till required.

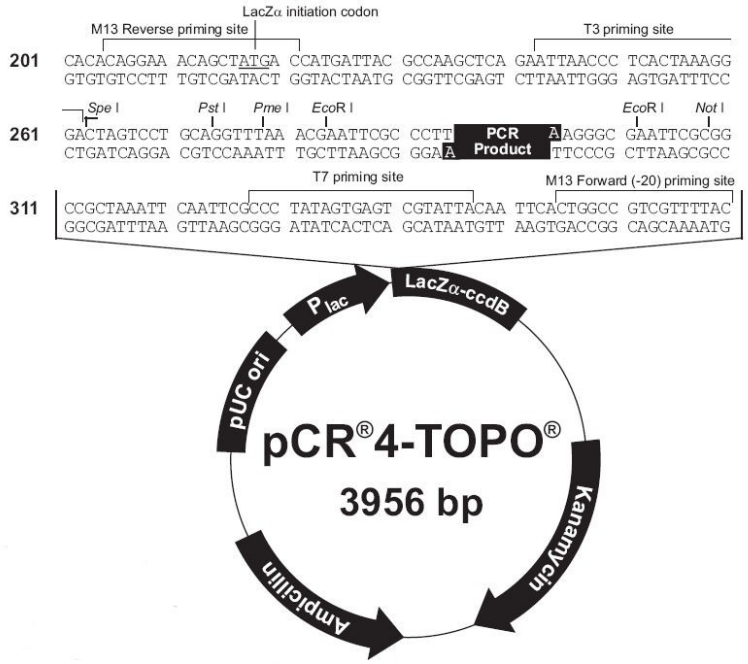


Figure 2.3. pCR 4-TOPO vector map. The pCR4-TOPO vector contains both ampicillin and kanamycin resistance markers, T7 and T3 promoters are also present for *in vitro* transcription. Figure from pCR TA-TOPO protocol Version R, April 2004 (Invitrogen).

2.3.1.1.1.Verification of the insert

In order to verify the orientation of the insert, plasmid DNA generated by Maxi Prep was digested using a combination of restriction enzymes, such as SalI and NotI, BglII and NotI (NEB, London, UK). These combinations were chosen as they cut once in the vector and once in the insert, providing distinctive different bands for either sense or antisense direction of a *myt1l* insert.

The restriction reaction was performed as follows: 2 µl of plasmid DNA (200 ng) was digested with 5U of each enzyme (BglII+NotI or SalI+NotI)(NEB, UK) in the presence 1x NEB restriction enzyme buffer 3 (NEB, UK) and 1x Bovine Serum Albumin (BSA) topped up with dH_20 to a final volume of 10 µl. The reaction was inclubated for 1h at 37 ^{o}C. The DNA digestions were separated by electrophoresis on a 1.5% w/v agarose gel (Sigma, UK) made up with 1X TBE buffer (0.0089 M Tris, 0.089 M Borate, 0.002 M EDTA; Merck, Germany) and supplemented with 2% of Ethidium Bromide (Electron, UK). 5 µl of the digested samples were mixed with 1 µl of 6X gel-loading buffer (Thermo Scientific, UK) and loaded into the gel wells. Additionally, one well was used for loading the 1 kb DNA ladder (250-10,000 bp Sigma, UK). Electrophoresis was

carried out at 120V for 1h. The DNA bands obtained from the gel were visualized using ultraviolet light.

2.3.1.1.2.Antisense RNA Probe Synthesis

After verification of the orientation of the insert, plasmid DNA was linearised with SalI restriction digest. The advised probe size was between 1 and 3 kb. Since the *myt1l* clone is ~4kb long, the SalI enzyme was chosen to create a partial clone, as it only cuts once in the insert (creating fragment of 1kb and 2.9kb in size), and it does not cut the vector.

The reaction was set as follows: 10 µg of plasmid vector was digested with 5U of SalI restriction enzyme(NEB, UK) in the presence 1x NEB restriction enzyme buffer 4 (NEB, UK) topped up with dH_20 to a final volume of 50 µl. The mix was incubated for 2h at 37 ^{o}C and then separated on a 1.5% agarose gel (120V for 1h) to check that the plasmid was fully linearised. The linearised DNA was then precipitated using 5 µl of 3M sodium acetate (NaOAc) and 125 µl of 100% biology-grade ethanol (Sigma, UK) and incubated overnight at -20 °C. The following day, the mix was centrifuged at 13,000 rpm for 30 minutes. The ethanol was removed and 300µl of ice cold 70% ethanol was added and then centrifuge at 13000 rpm for 15 minutes. The ethanol was removed and the DNA pellet was air dried. The DNA was then resuspended in 50 µl of dH_20, 1 µl of the reaction was checked on a 1.5% agarose gel (120V for 1h), and the DNA template was stored at -20°C. 1 µg of purified, linearised DNA template was transcribed and labelled with digoxigenin using 4 µl of 5 x transcription buffer (Promega, UK), 2 µl 10 x DIG RNA labelling mix (10mM each of ATP,CTP,GTP, 6.5mM UTP, 3.5mM DIG-11-UTP, pH 7.5 (Roche) 1 µl of T3 RNA polymerase (Promega), 2µl of 0.1M DTT 1 µl of RNaseOUT Recombinant RNase Inhibitor (Invitrogen) and the volume was made up to 20 µl with nuclease-free water. The reaction mix was incubated at 37°C for 2 hours. DNA was digested by adding 1 µl of RNAse-free DNAseI (Roche) for 30 mins at 37°C. Synthesis was stopped by adding 1 µl of 0.5 M EDTA and the mix was purified using QIAquick PCR Purification columns. 1 µl was checked by electrophoresis on the 1.5% agarose gel (120V for 1h) for the probe quality. A 120 µl of Hybridisation Mix (HM) (Table 2.1) was added to the probe and the mix was stored at -20°C.

2.3.1.2. Whole Mount in situ Hybridisation Protocol

The solutions used for *in situ* hybridisation are listed in Table 2.1. Embryos were fixed at the required stages (18somite stage, 24hpf, 48hpf, 72hpf and 120hpf) in 4% paraformaldehyde (PFA) overnight at 4°C. For longer term storage embryos were dehydrated in 100 % methanol. The PFA was washed out 3 times by soaking in PBS with 0.1% Tween20 (PBT) for 5 minutes before dechorionation. The embryos were manually dechorionated and dehydrated through a series of incubations in 25%, 50%, 75% and 100% methanol in PBST for 5 minutes in each step. 100% methanol was added to cover the embryos. The embryos were then stored at -20°C for future use.

Table 2.1. Solutions for *in situ* hybridisation.

Solution	Ingredients
20x Saline Sodium Citrate (SSC)	3M NaCl, 0.3 M Citric acid trisodium salt, pH7
Maleic acid buffer(MAB)	100 mM maleic acid, 150 mM NaCl, adjusted to pH 7.5 with NaOH
Hybridisation Mix (HM)	50% formamide, 5 x SSC, 0.1% Tween20, 50µg/ml Heparin (Sigma), 500 µg/ml RNAse Free tRNA Torula (Sigma)
Staining buffer (NTMT)	100 mM tris HCl pH 9.5, 50 mM MgCl, 100 mM NaCl, 1% Tween20
PBT	1 x PBS (Dulbecco's Phosphate Buffered Saline Modified, without CaCl2/MgCl2, Sigma), 0.1% Tween20

Whole mount ISH was performed as previously published (C. Thisse & Thisse, 2008) with some modifications as described. All steps took place in 1.5ml tubes. Figure 2.4 outlines the stages and timing involved in the ISH procedure.

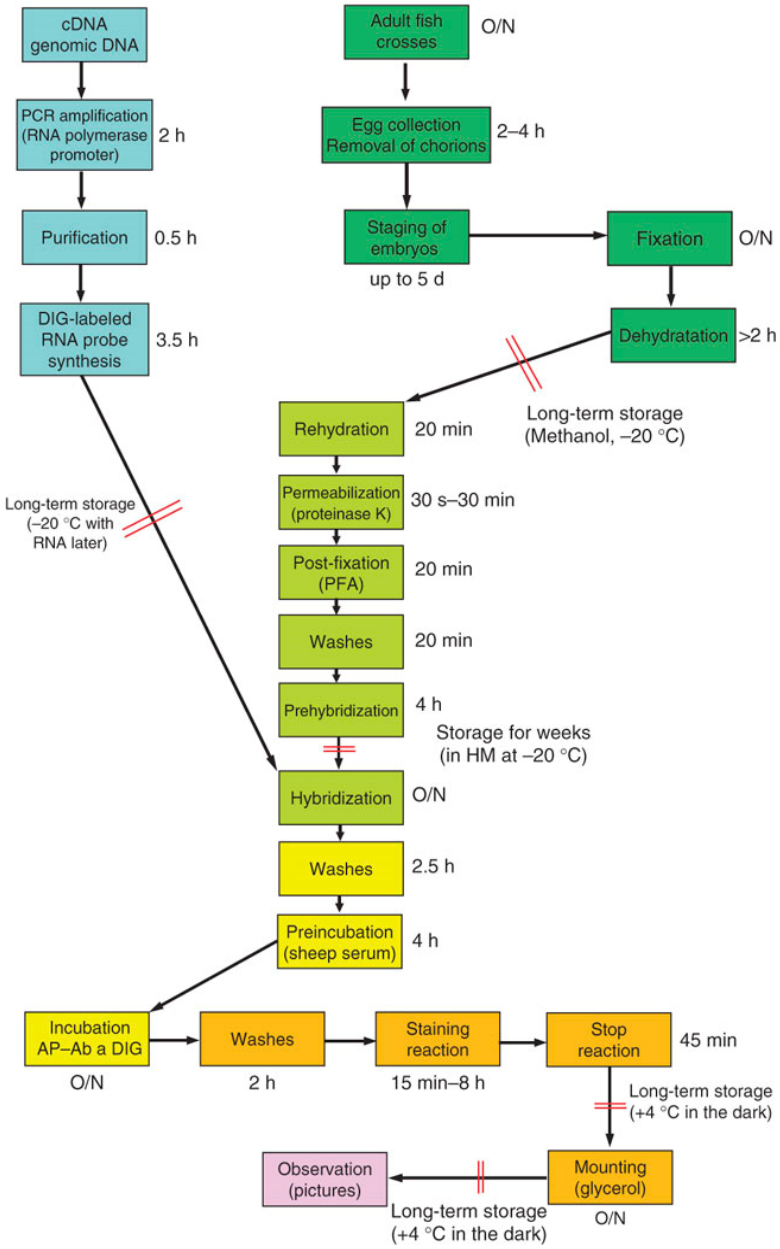


Figure 2.4 Figure outlining the stages and timing involved in the *in situ* hybridasation (ISH) protocol. The different steps of the ISH are indicated in boxes linked by arrows and coloured: blue for steps corresponding to synthesis of the probe, dark green for preparation of embryos, light green for day 1 of *in situ,* yellow for day 2, orange for day 3 and pink for the final step. Time required for each step is indicated near the box (O/N: overnight). Pause points are indicated in the diagram by two red lines across arrows. (Picture from (C. Thisse & Thisse, 2008))

Day 1: Embryos were gradually rehydrated from methanol to 100% PBT by a series of 5 minutes incubations in 75%, 50% and 25% methanol in PBT. Embryos were then washed once with 0.5ml of 100% PBT for 5 min. Proteinase K (0.3ml of 50µg/ml) treatment for embryos older than 18SS (somite stage) was performed to ensure permeabilisation of embryos and to permit an access of RNA probe. The length of this treatment, which depended on the

developmental stage, was as follows: 1 min for 24hpf, 2 min for 32hpf, 5min for 36hpf, 8min for 48hpf, 10 min for 72hpf, 12 min for 96hpf and 14 min for 120hpf. Proteinase K treatment was stopped by replacing it with 0.5ml of HM. Embryos in HM were prehybridised for at least 4h at 65°C horizontally on a heatblock. DIG-labelled probe was diluted 1:100 in HM and pre-warmed for 5 minutes at 65°C. Pre-hybridisation solution was removed and embryos were incubated in 0.3ml of prewarmed diluted probe overnight at 65°C on a heatblock.

Day 2: Hybridisation solution-containing probe was removed and embryos were washed with prewarmed 0.5ml HM solution for 30min. To remove non-specifically hybridised probes serial washes with SSC were carried out as follows: 1x 10min with 2XSSC (1ml) at 65°C, 2x 30 min with 0.2X SSC(1ml) at 65°C followed by 1x 5min in containing 0.1% Tween20 (1ml) (MABT). Embryos were then incubated in blocking buffer containing MAB supplemented with 1% Blocking reagent (Roche) for 3-4h at room temperature on the shaker. The blocking solution was replaced with 0.3ml of blocking buffer containing 1:5000 anti-DIG antibody and incubated at 4°C overnight while shaking.

Day 3: Antibody solution was discarded and embryos were washed in 1ml MABT 0.1% at least 8 times, 15-20 min each, equilibrated for at least 20min with NTMT staining buffer and transferred to 6 or 8-well plated. Colour development was performed in the dark at room temperature in 0.3ml of the staining solution (3.375 µl NBT (Sigma)(100mg/ml), 3.5 µl BCIP (Sigma)(50mg/ml) in 1ml NTMT). Embryos were observed every 10-15 min under the dissecting microscope to monitor colour development then theplate was placed in the dark to avoid overexposure to light. To stop the colour development, embryos were washed 3 times for 5 mins with PBT solution and incubated in 4% PFA for 20 minutes at room temperature. The PFA was removed by three 5 mins PBT washes. The embryos were stored in 70%Glycerol/PBT at 4°C.

2.3.2. Visualisation and Image Capture

Embryos were mounted on slides and cover slips in order to visualise the probe staining and record results, following the procedure below. To mount the embryos four cover slips of thickness 1.5mm were stacked together to form a bridge. Stained embryos were placed in a drop of glycerol in the middle of a microscope slide between bridges. The bridge and the embryo in the drop of glycerol are covered with a large 24 x 40 mm coverslip. Due to the

viscosity of the glycerol, the embryo could be rolled and observed in different orientations by gently moving the 24 x 40 mm coverslip.

2.3.3. RNA extraction from whole zebrafish embryos.

RNA is relatively unstable and easily prone to cleavage by RNAses, therefore it is of utmost importance to clean the work area by spraying the work surfaces with a decontaminant reagent (RNaseZAP) before beginning the procedure. To attain a sufficient amount of RNA, embryos of each stage were divided into groups of 50 or 20 (depending on the stage, 50 embryos for 18SS, 25 embryos for 24hpf and 20 embryos per 48hpf and 72hpf) transferred to separate 1.5ml tubes and excess water was removed with a pipette. The embryos were then frozen by placing the tubes on dry ice and transferred to -80°C freezer until ready for extraction. Before starting the extraction PhaseLock heavy tubes (FisherScientific) were prespun at 1500 RPM for 30sec.

One pool of samples was removed from -80°C, placed on ice and immediately 250 µl of Trizol reagent (Invitrogen) was added to the tube. Embryos were lysed and homogenised with a disposable tissue grinder (Sigma). When cells were sufficiently homogenised a further 250 µl of Trizol reagent was added. Samples were then transferred to prespun PhaseLock tubes (FisherScientific), mixed by inverting the tube and incubated at room temperature for 5 min. 100 µl of chloroform was added to the samples and mixed vigorously for 20s before further incubation for 2min. Samples were then centrifuged at 13,000 x g for 15 minutes at 4°C. The colourless upper phase of the sample was transferred into an RNase free 1.5ml tube, using a 1ml pipette and being careful not to transfer any of the interphase layer, and 250 µl of isopropanol was added to precipitate the RNA. Samples were allowed to sit at room temperature for 10 min and then centrifuged at 13,000 x g for 10 min at 4°C. RNA formed a gel-like pellet on the bottom of the tube. Without disturbing the pellet, isopropanol was removed with a pipette and the pellet was washed with 0.5ml of 100% ethanol (for molecular biology use, Sigma). Without disturbing the pellet samples were mixed by gentle inversion followed by centrifugation at 13,000 x g for 5 min at 4°C. After centrifugation, the ethanol was removed, and the pellet was allowed to air dry while inverted for 10min. The pellet was resuspended by adding 100 µl DEPC-treated (RNAse-free water)(Sigma).

2.3.4. RNA clean-up.

Qiagen RNease Mini-Kit was used for total RNA clean-up and manual instructions were followed. Briefly, to each sample prepared as described in section 2.3.2, 350 µl of RLT buffer containing β-mercaptoethanol (β-ME) (10 µl of β-ME per 100ml of Buffer RLT) was added and samples were mixed well by pipetting then centrifuged for 3 minutes at 13,000 x g. Supernatant was carefully removed and transferred to a new 1.5ml tube and used in subsequent steps. 350 µl of 70% ethanol was added to the samples and mixed well by pipetting. The sample was then transferred to an RNeasy spin column placed in a 2 ml collection tube and centrifuged for 15 seconds at 8,000 x *g* and the flow through was discarded. 350 µl of Buffer RW1 was added to the RNeasy spin column and centrifuged for 15 seconds at 8,000 x *g*. The flow-through was discarded and the RNeasy column was re-used. DNAse treatment with Qiagen RNAse-free DNAse kit was performed as follows. Per sample, 10 µl of DNAse I stock solution (DNAse I powder dissolved using 550 µl of RNase-free water) was added to 70 µl of buffer RDD and the mix was added to each sample and incubated for 15 min at room temperature. Following DNAse treatment, 350 µl of buffer RW1 was added to the spin column and samples were centrifuged at 8000 x g for 1 min. 500 µl of Buffer RPE (containing 70% ethanol) was added to the RNeasy column and centrifuged for 15 seconds at 8000 x *g*. The flow through was discarded and a second aliquot of 500 µl of Buffer RPE was added to the RNeasy spin column followed by centrifugation at 8000 x *g* for 2 minutes. The RNeasy spin column was then placed in a new collection tube and 30 µl of RNA-free water was added directly to the spin column membrane and centrifuged for 1 minute at 8000 x *g* to elute the RNA. RNA concentrations were determined using the Nanodrop-1000 spectrophotometer (Thermo Scientific, UK) and the quality checked by referencing the ratio of absorptions at 260/280nm in order to detect protein contamination. Samples were then stored at -80°C freezer.

2.3.5. Mouse RNA

RNA samples extracted from CD1 mouse brains at embryonic day 10 (E10), E14, E18, and at postnatal (P) stages 1 week, 1 month, or 6 months were obtained from AMS Biotechnology, Abingdon, UK. RNA samples from the E10 and E14 stages were pooled from 5 and 3 brains, respectively. Triplicate samples from E18, P1week and P1month prefrontal cortex were

obtained from independent brains, and the P6month prefrontal cortex RNA sample was derived from a single mouse brain.

2.3.6. cDNA Synthesis

cDNA synthesis was performed by reverse transcription of RNA from both, zebrafish and mouse. First strand cDNA synthesis was carried out on total RNA using the Superscript™ III synthesis system (Invitrogen, Paisley, UK). 1 µg of total RNA was added to 1 µl of 10mM dNTP mix and 1 µl of Oligo$(dT)_{18\text{-}20}$ (0.5 µg/µl) and the volume made up to 10 µl with RNAse-free water. Each sample was incubated for 65°C for 5 minutes and then placed directly on ice. Next, to each sample 2µl of 10x RT buffer, 4µl of 25mM $MgCl_2$, 2µl of 0.1M DTT, 1µl of RNaseOUT™ Recombinant RNase Inhibitor and 1µl of Superscript™ III RT (2000U/µl) was added. The samples were incubated at 50°C for 50 minutes and the reaction was terminated by 5 minutes at 85°C. 1µl of RNase H was added to each sample and then samples were incubated at 37°C for 30 minutes followed by sample dilution with distilled water, making 100 µl of each sample at a concentration of 10ng/µl.

2.3.7. Quantitative polymerase chain reaction (qPCR)

Real-time quantitative PCR (polymerase chain reaction, q-rtPCR, qPCR) is a technique used for amplifying DNA and monitoring the progress of the reaction in the 'real time'. To allow the progress of a PCR to be monitored a fluorescent marker (e.g. SYBR green) is used which binds to the DNA. Thus, as the number of gene copies increases during the reaction, the fluorescence increases as well. This is beneficial because the efficiency and rate of the reaction can be measured.

cDNA samples were amplified using an ABI Prism 7900HT sequence detection system in a final volume of 20 µl. 10 µl of SYBR© Green PCR Master Mix (Applied Biosystems), 0.5µM of primer F and 0.5µM of primer R were added to 4µl of cDNA template and topped up with RNase-free water. The qPCR reaction was performed in triplicates. The amplification procedure consisted of an initial enzyme activation step of 95°C for 15 min followed by 40 cycles of 30 sec at 95°C and 30 seconds at 59°C. Fluorogenic data was collected at the 59°C stage and the data were analysed using SDS software v 2.3 (Applied Biosystems). The PCR reaction products were

evaluated by a melting curve analysis. Sequences of the primers used can be found in Table 2.2. *Gapdh* was used as an internal control for mouse RNA, and *β-actin* was used as a housekeeping gene for zebrafish samples.

Table 2.2. Primer sequences used for establishing expression levels of myt1l by qPCR.

Primer		Sequence
zf-βactin	Forward	5' – GCAGAAGGAGATCACATCCCTGGC – 3'
	Reverse	5' – CATTGCCGTCACCTTCACCGTTC – 3'
zf-myt1l	Forward	5' – TGATGAGCACGATGAAGAGG – 3'
	Reverse	5' – TTCGTGTTGTTGTTGTTGTTG – 3'
m-Gapdh	Forward	5' – TGTTCCTACCCCCAATGTGT – 3'
	Reverse	5' – CCTGCTTCACCACCTTCTTG – 3'
m-Myt1l	Forward	5' – TGGTGACGATGTAGAAGAGGA – 3'
	Reverse	5' – TCCTTGTCTGTGTCCTGCAT – 3'

Firstly, mRNA levels were first normalised to that of housekeeping gene to generate ΔCt values ($\Delta Ct = Ct_{target} - Ct_{housekeeping\ gene}$) at each developmental stage. Changes in expression at each stage were relative to E10 for which ΔΔCt ($\Delta Ct - \Delta Ct_{E10}$) were calculated. ΔΔCt was then converted to relative fold expression using the formula $2^{-\Delta\Delta Ct}$. Data are presented as group means and error bars show the standard error. All RT-qPCR data were analysed by one-way ANOVA, followed by Tukey *post-hoc* analysis with $p \leq 0.05$. Statistical analysis was performed comparing expression of triplicates at the E18, P1week andP1month stages to that of E10.

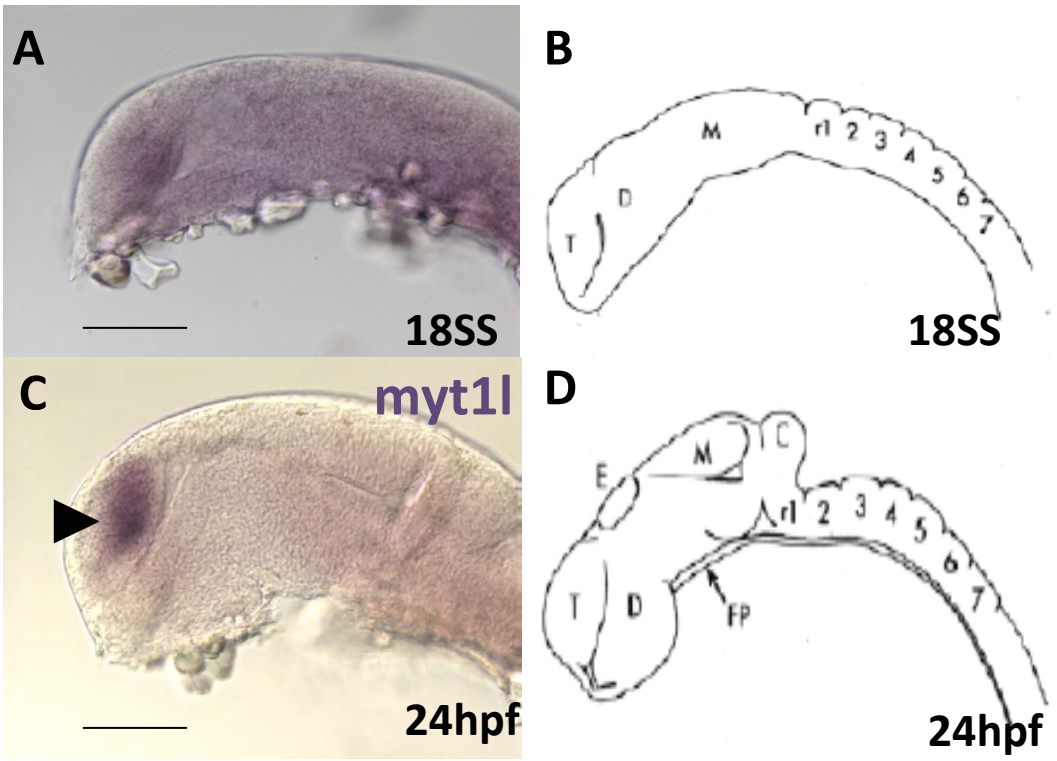
2.4. Results

2.4.1. myt1l tissue distribution during zebrafish development

Our first action was to interrogate the Ensembl database to find MYT1L orthologs. We have found that myt1la is 68% homologous to human MYT1L and mouse myt1l. The *myt1l* gene is located on chromosome 20, with 4 known transcripts, the longest spanning 6347bp. So far no studies have examined the developmental distribution of myt1l during zebrafish development. Its expression patterns were investigated by whole-mount *in situ* hybridisation (ISH). ISH was performed on 18SS, 24hpf, 32hpf, 36hpf, 48hpf, 72hpf and 120hpf with a probe developed from PCR product as described in section 2.3.1.

All the figures shown below represent only the developing nervous system region and the zebrafish head, as there was no visible staining anywhere else in the zebrafish body (data not shown).

As shown in Figure 2.5A, myt1l mRNA expression was not found in the 18 somite-stage embryo. At this stage we did not detect a presence of the labelled probe, suggesting that if present, levels of myt1l were low and not detected by the probe.



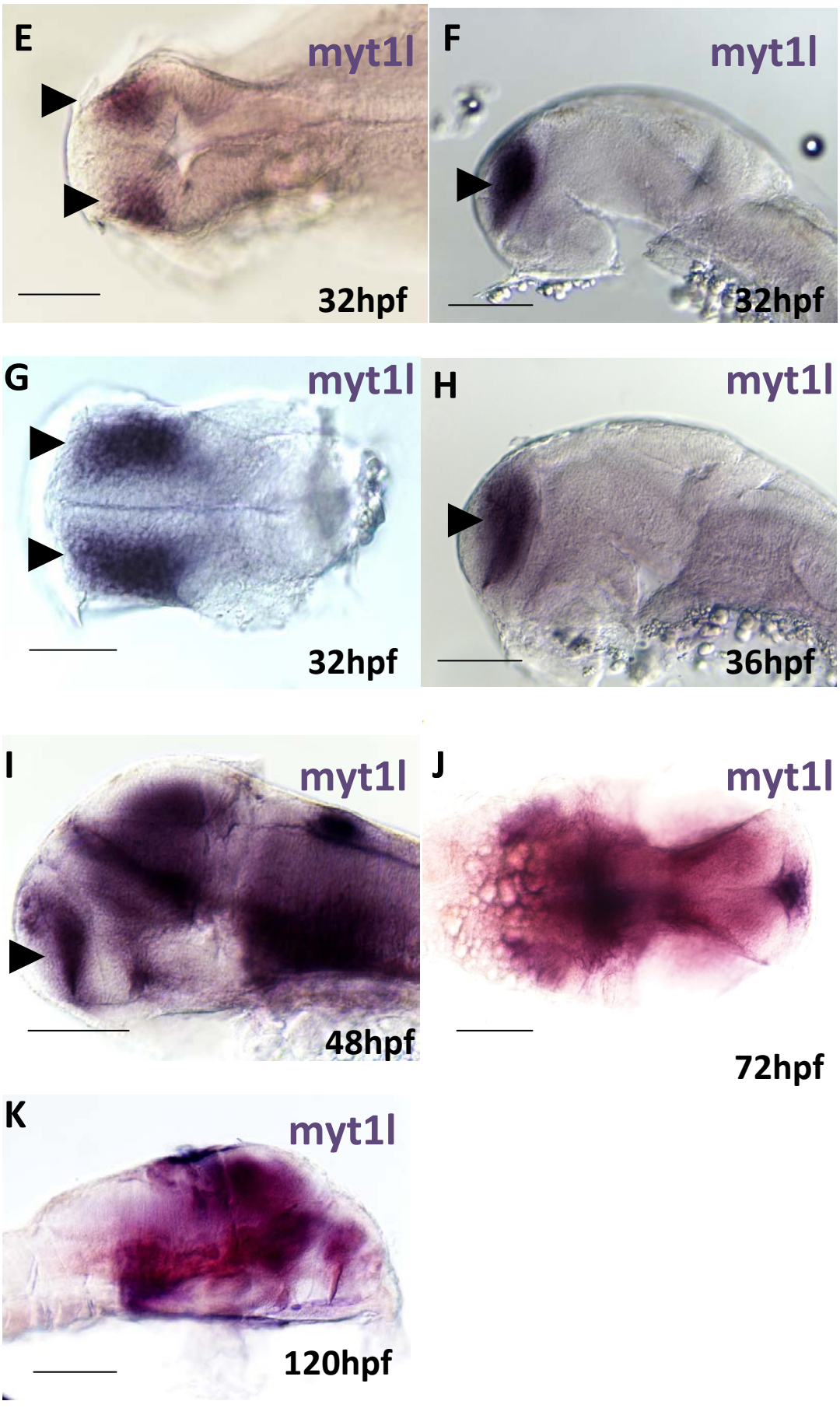

E
myt1l
32hpf
F
myt1l
32hpf
G
myt1l
32hpf
H
myt1l
36hpf
I
myt1l
48hpf
J
myt1l
72hpf
K
myt1l
120hpf

Figure 2.5. **Whole mount *in situ* hybridisation of *myt1l* expression from 18ss-120hpf.** Figures A,C,F,G,H,I, K are lateral views of zebrafish embryo head, D,E,J are frontal views of a zebrafish embryos displaying bilateral *myt1l* expression.Stages are indicated in the right bottom corner of each panel **A**- 18 somite stage showing no detectable expression. Figure **C** and **E** are 24hpf clear expression in the forebrain area as indicated by purple stain. **E-G** view of 32hpf zebrafish head displaying bilateral *myt1l* mRNA expression in the forebrain region, **H** show strong expression of *myt1l* in the 36hpf embryo, restricted to the forebrain. **I-K** represent views of widespread *myt1l* expression in the embryonic brains at 48hpf, 72hpf and 120hpf respectively. Figure B represents schematic of the brain segmentation at 18SS, the telencephalon (T), diencephalon (D), mesencephalon (M), and about seven hindbrain rhombomeres (r1-r7). **D** schematic of brain development at 24hpf, the epiphysis (E) is present in the midline of diencephalic roof. The dorsal midbrain, or tectum (M), is partitioned from the ventral midbrain. The cerebellum (C) is at the hindbrain/midbrain boundary region. The floor plate (FP) extends in the ventral midline up to, but not including the forebrain (figure from(Charles B. Kimmel et al., 1995). Scale bars are 40 μm. Arrowheads indicate telencephalon

We detected the first expression at 24hpf (purple stain, Figure 2.5C), in the telencephalon, which is the anterior subdivision of the embryonic forebrain or the corresponding part of the adult forebrain that includes the cerebral hemispheres and associated structures. *myt1l* expression was not observed in any other part of the embryo, at any studied stage (data not shown), which confirms that the expression of *myt1l* is brain specific. Continuous expression of *myt1l* was clearly visible at 32hpf and 36hpf stages (Figure 2.5 F-G and H), its expression being still localised in the telencephalon with no staining observed in midbrain/hindbrain structures. By the second day of development (48hpf) *myt1l* expression became much more widespread throughout the CNS (Figure 2.5I). As expression domains were relatively large, it became difficult to precisely define myt1l expressing brain regions in whole mount preparations. The process of identifying specific brain regions in which myt1l was expressed became even more challenging for later stages of embryonic development, as the staining was very robust across the whole brain area. The pattern of expression at 72hpf and 120hpf was similar to that of 48hpf, however during this additional 24hr period some brain regions exhibited progressively more intense and expanded expression (Figure 2.5 J-K). The staining at these stages was not improved by various experimental alterations, such as extended periods of Prokinase K treatment to allow the fluorescent probe better penetrate the embryo.

These results show that myt1l mRNA transcripts are brain specific, and its presence was not detected during early embryonic growth but it is strongly expressed at the later stages during brain development.

2.4.2. qPCR analysis to measure myt1l mRNA levels in zebrafish embryos.

As mentioned above, we have found brain specific myt1l expression during zebrafish embryonic development starting between 18SS and 24hpf. To examine the magnitude of change in expression, myt1l mRNA levels at various stages of zebrafish development were measured by qPCR. To accurately quantify gene expression, the expression levels of myt1l were normalised to the expression level of β-actin producing ΔCt. Since it is the first insight into myt1l mRNA levels distribution in zebrafish, examining ΔCt values provides more information about its expression in all samples, as it does require to relate changes in gene expression to a reference sample. We have found that myt1l is detected by qPCR at 18SS, however its levels at this stage are low (Figure 2.6). With each developmental stage expression gradually increased, which is consistent with ISH findings.

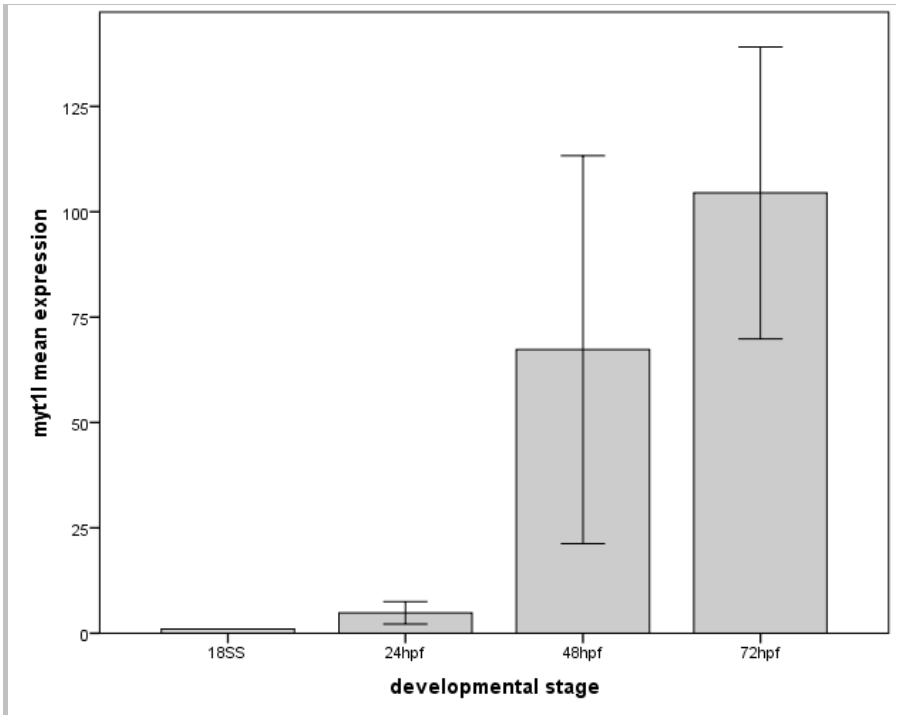


Figure 2.6. qPCR analysis of *myt1l* expression in zebrafish embryos. Highest expression (100-fold change) is reported from 3 day embryos. Results from 24hpf, 48hpf and 72hpf were calculated relative to the expression at 18SS stage. Data is presented as mean values of three independent experiments, error bars represent +/- 1SE of the mean.

The results did not reach the significance level, when analysed by one-way ANOVA ($F_{(3,8)}$= 3.04, p=0.093). It is interesting to note that although the samples used for qPCR were whole embryo-derived and overall brain-specific RNAs within these sample was diluted, we were still able to detect expression changes across different time points.

By examining the dissociation curves, the specificity of the qPCR products was assessed. Inspection of the dissociation curves for zebrafish *myt1la e*xpression in RT-qPCR experiments revealed a single, specific amplicon for each measurement, confirming primer specificity (Figure 2.7).

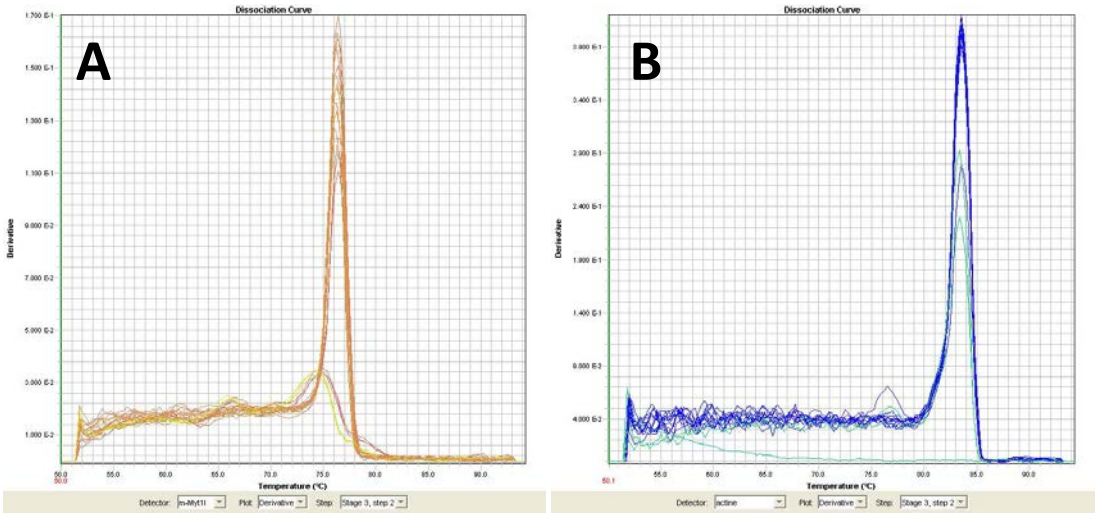


Figure 2.7. Confirmation of the specificity of m*yt1l* RT-qPCR primers. Dissociation curves from qPCR analysis of *myt1l* (A) and *β-actin* (B) expression in zebrafish embryos indicate a single, specific amplicon (i.e. one PCR product per primer pair). The additional peaks visible on both figures were generated from negative controls only (i.e. water).

To further validate our findings and to check for possible primer dimers, we run 5µl of each qPCR sample on 2% agarose gel with expected amplicons of around 200bp for myt1l and around 400bp for β-actin (Figure 2.8).

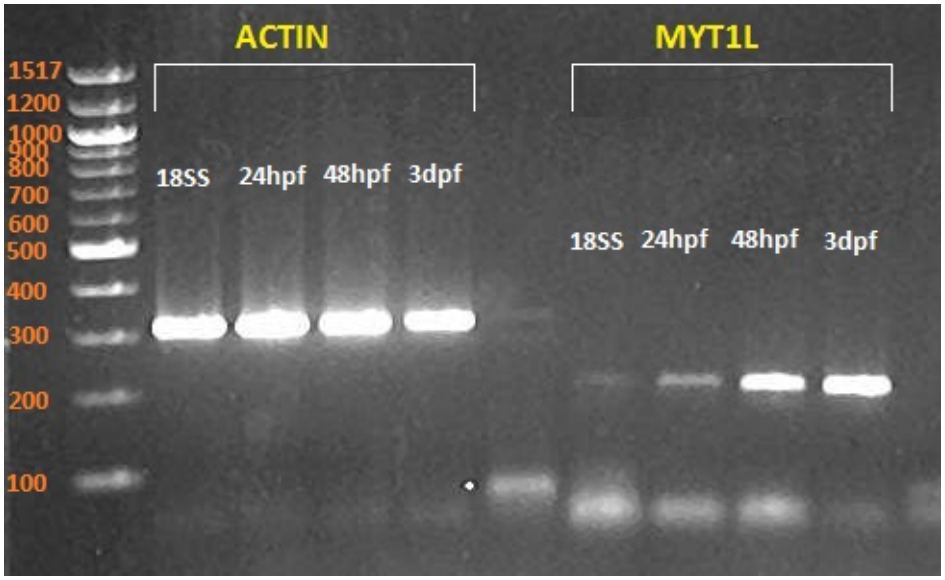


Figure 2.8. Agarose gel electrophoresis of qPCR samples. Lanes 1-4 represent qPCR samples with actin (18SS, 24hpf, 48hpf and 72hpf respectively), lanes 6-9 correspond to myt1l amplified qPCR products (18SS, 24hpf, 48hpf and 72hpf respectively), lane 5 was left empty and lane L was the 1kb DNA ladder used as a marker.

Gel electrophoresis of *myt1l* qPCR samples demonstrated a Myt1l transcript with band size of approximately 200bp, corresponding with the predicted size of 198bp (Figure 2.8). Furthermore, clear expression patterns were visualised on the gel. At 18SS stage the band is very faint and hardly visible, in the 24hpf embryos, myt1l is represented as a thin and grey band, demonstrating the higher levels of *myt1l*. Presence of the qPCR product is evidently shown at 48hpf with a very strong and thick band on the gel. The expression is even higher at 3dpf with an even more intense band being observed on the gel. These results confirm that myt1l mRNA level changes during embryonic development, with very little gene product present at 18SS stage and progressively increasing during development until 3dpf. We observed extra bands in the lanes containing myt1l samples which could suggest a slight gel contamination or possible a non-specific amplification. Small faint bands below the primary amplicons are likely to be primer-dimers or cDNA related amplicons and may be dismissed provided the dissociation curve displays only a single peak for each product.

2.4.3. Myt1l expression in mouse brain samples.

The validation of our findings and a further assessment of the changes of Myt1l transcript levels in the vertebrate brain were made by investigating changes in expression of this transcription factor in the mouse brain during embryonic and postnatal development. To

measure expression levels we performed qPCR analysis of *Myt1l* expression and the results were first normalised to that of *Gapdh* at each developmental stage. Inspection of the dissociation curves for *Myt1l* expression in RT-qPCR experiments revealed a single, specific amplicon for each measurement (Figure 2.9), which indicates that the primers only amplified the region for which they were designed.

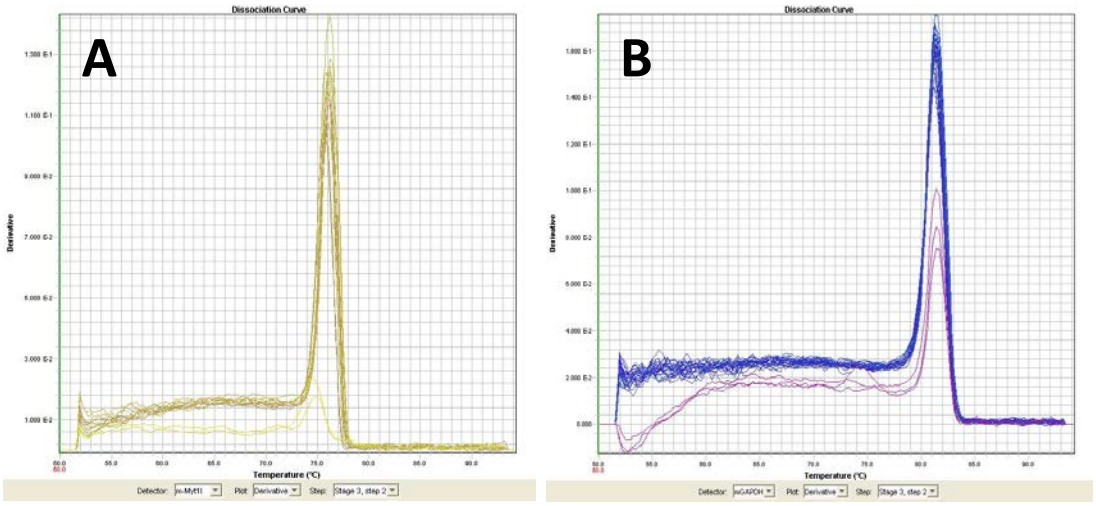


Figure 2.9. Confirmation of the specificity of *Myt1l* RT-qPCR primers. Dissociation curves from qPCR analysis of *Myt1l* (A) and *Gapdh* (B) mRNA expression in in mouse brains indicate a single, specific amplicon (i.e. one PCR product per primer pair). The additional peaks visible on both figures were generated from negative controls only (i.e. water).

We observed an increase in *Myt1l* expression at each developmental stage. Highest expression of *Myt1l* is observed at E18 (Figure 2.10) which is just before the litter is born, and remained elevated at 1week and 1 month after birth, which corresponds to adolescence in mice . *Myt1l* expression gradually decreases at each postnatal stage. Results showed a 500-times increase in *Myt1l* expression at 6 months post-natally although lack of replicates prevented determination of statistical significance for this stage. A one-way ANOVA revealed a significant gene dosage effect in the mouse brain samples ($F_{(3,8)}$= 13.175, p=0.002). Post-hoc analysis using Tukey shown that the change was significant between E10 and E18 (p=0.003) and between E10 and 1 week after birth (p=0.004).

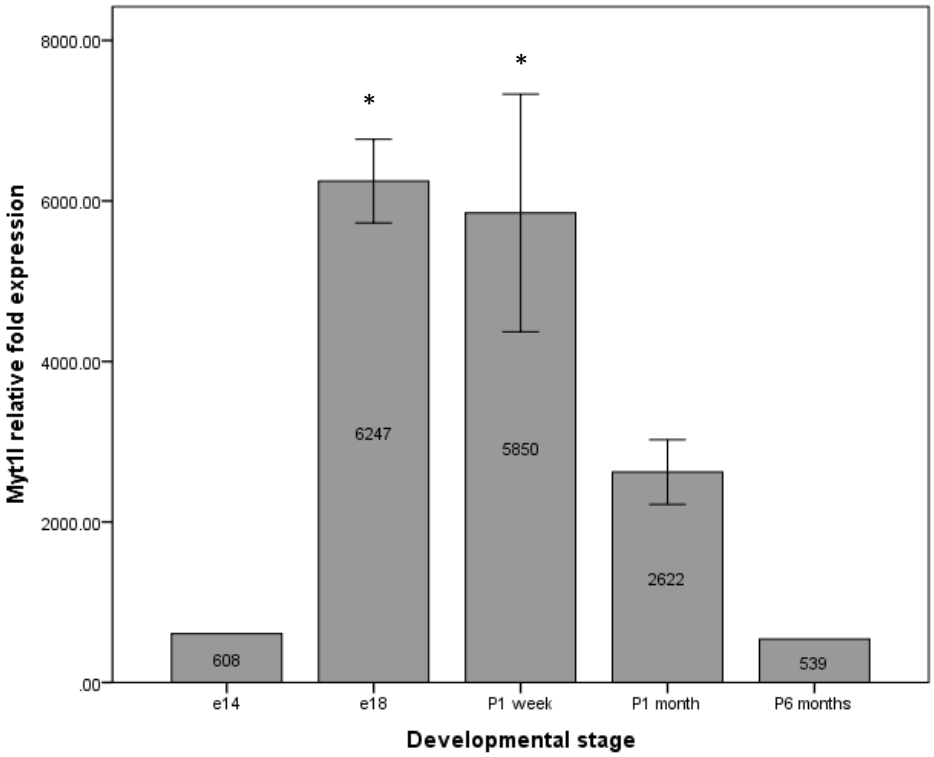


Figure 2.10. Graph representing stage specific expression of *Myt1l* during mouse brain development. The highest expression was recorded at E(embryonic day)18. Data from E14, E18, P1 week, P1 month and P6 months were calculated relative to the expression in the brain at E10. Statistical analysis (ANOVA) compared expression at E10, E18, 1 week and 1 month post natally. Tukey post hoc analyses showed significant increase in expression at E18 (*-p=0.003) and at 1 week after birth compared to E10 (*-p=0.035). Error bars represent +/- 1SE of the mean.

2.5. Discussion and study limitation

This study aimed at characterising the developmental distribution of *Myt1l* in vertebrates, by two two model organisms – zebrafish and mouse. To visualise the localisation of zebrafish *myt1l, in situ* hybridisation experiments were performed. We found that *myt1l* is expressed specifically in the zebrafish brain, starting between 18SS and 24hpf, with expression remains high throughout the embryogenesis. At the end of the gastrulation period (10hpf), the first onset of neurogenesis in the neural plate becomes apparent. At 24hpf this gene is clearly showing in the telencephalon, at the time when neuronal progenitor cells start proliferating in this brain region (Schmidt, Strahle, & Scholpp, 2013), suggesting involvement of *myt1l* in neurogenesis and its putative role as a developmental regulator. Our findings are consistent with recent genome-wide analysis study, by Armant et al.(Armant et al., 2013), where it was found that *myt1la* was one of 16 genes that were only expressed in the telencephalon at 24hpf, alongside patterning and differentiation genes such as *emx*1, *emx*3, *tbr*1*a* and

neurod6b. The authors of this study suggested that myt1l alongside other identified genes is a prime candidate for functional studies due to their single tissue specificity. Within the first 24h of fish embryogenesis, a precisely organised and simple network of neurons develops. These neurons form a foundation for central and peripheral axonal pathways (C. B. Kimmel, Hatta, & Eisen, 1991). At 24h, neurogenesis has already started and continues; regionalisation starts in the pallium (septal formation) and subpallium (striatum). Some neurons begin to connect to their targets in the diencephalon. At this stage in the telencephalon, neurons of the dorso-rostral cluster (drc) extend axons to form the supra-optic tract (SOT) and the anterior commissure (AC)(Hjorth & Key, 2002). The zebrafish telencephalon is composed of the pallium and the subpallium, which is the teleost analog of cortico-basal-ganglia circuits in mammals (Rink & Wullimann, 2002). In rodents and non-human primates, the formation of new neurons is particularly evident in two regions, the subventricular zone (SVZ) and the subgranular zone (SGZ) of the telencephalon (Schmidt et al., 2013). Exclusive presence of myt1l within this brain region from 24hpf till 36hpf supports the suggestion that this gene is involved in neurogenesis.

As development proceeds, Myt1l transcript becomes detectable in many other areas of the brain and it becomes difficult to definitively define *myt1l* expressing brain regions in whole mount preparations. By 48 hpf and in the early larval period (from 72 hpf), expression of *myt1l* was observed throughout the head, but it was not detected anywhere else in the fish body. It is generally accepted that by 48-72h second neurogenesis occurs with the beginning of the hatching period (Charles B. Kimmel et al., 1995). This stage marks the origin and development of neurons building the main mass of the later brain. The 2-dpf zebrafish brain is unique in that the key markers of neurogenesis were found to be expressed in a pattern which is to be expected for the beginning of massive overall secondary neurogenesis and represents a first stage of zebrafish brain development comparable with what is described in the mouse brain (Mueller & Wullimann, 2003).

To confirm the expression patterns of *myt1l* during zebrafish development, we performed qPCR across developmental stages. Similarly to ISH, we concluded that *myt1l* is highly expressed at 24hpf and its levels remain elevated until 3dpf. Our study is the first to provide an insight into myt1l expression across different embryonic stages during zebrafish development. Importantly, strong expression of *myt1l* was seen in the developing brain from 24 hpf to 3 dpf, suggesting its possible role in zebrafish neural development.

To further elucidate the role of this gene and to examine if its expression followed similar pattern in higher vertebrates, RNA extracts from mouse brains at various developmental time

points were analysed. Similarly to zebrafish, we found that *Myt1l* was not expressed early during the development (up to E13.5), and its levels were the highest just before birth (E18). We found that Myt1l mRNA levels remained elevated but gradually decreasing up to 6 months post-natally (last measured stage). This can suggest that although still highly expressed, Myt1l most robust role is before birth, when vertebrate neurogenesis is most active. Interestingly, we noted high levels (around 600-fold increase) of *Myt1l* at 6-month postnatally which corresponds to mature adult in human. These facts can imply that MYT1L might not only play a role during embryonic brain development but it can also have a functional part in adult neurogenesis, like maintaining synaptic functions and connections throughout the life. Discovering that *Myt1l* is highly expressed later in life can suggest its role as a possible neuroprotective agent or that it may be a factor required to maintain neuronal functions.

Most interestingly we can conclude that *Myt1l* expression patterns are conserved throughout vertebrate development, and that its expression is restricted to brain. These suggest a pivotal role for this gene in vertebrate neurodevelopment.

2.5.1. Limitations and future directions

The first and foremost advantage of our study, which that is at the same time a limitation, is the choice of zebrafish as a model organism. Though a vertebrate, the zebrafish is a non-mammalian species, implying physiology and pathology to be more evolutionarily distant from human than mice, the most commonly utilised model organism.

One limitation comes from in situ hybridasation staining. Using this method, we were only able to precisely visualise region-specific *myt1l* transcripts up to 48h of embryonic development. This could have been overcomed by performing tissue sections of fish brain. It would have given us more insight into the specific organisation of *myt1l* expression and its possible role in neurogenesis. Direct detection of myt1l protein can be achieved through immunohistochemical methods. Immunostaining is a powerful method for detecting the presence and the localisation of an endogenous protein. Analyses of the expression and the localisation of proteins are crucial to determine the function of genes at the tissue and organ levels mainly by immunohistochemical and live-imaging analyses. It is important to validate the expression and sub-cellular localisation of endogenous proteins. This is best analysed by immuno-histochemistry in sections or whole mount preparations. To determine which type of cells express myt1l the brain sections could have been co-immunolabelled with myt1l and

specific cellular markers (i.e. for GABA-ergic neurons, glutaminergic neurons, serotonergic neurons). Additionally, it would be interesting to investigate myt1l expression patterns and levels in later stages of fish development. It is known that during adulthood the brain remains plastic and behavioural changes like memory and learning can influence synaptic function and organisation (Spedding, Jay, e Silva, & Perret, 2005). Our mice data suggest that Myt1l might play a role during adulthood, thus it would be interesting to further examine the zebrafish model and potentially identify brain regions associated with myt1l expression. This information would provide us with valuable clues as to the function of this transcriptional factor. As for future directions, to further our knowledge and understanding of *Myt1l* role in vertebrate brain development, its down-regulation could be investigated. Interference with this gene function *in vivo* can produce invaluable data as to its role in fish brain development. There are several widely implemented techniques in both, zebrafish and mouse models that could potentially be implemented to study *Myt1l* function. For example, use of morpholino-modified oligonucleotides (morpholinos) is an effective knock-down technology widely used in the zebrafish. Morpholinos are synthetic complementary antisense oligonucleotides that can either block translation (ATG morpholinos) or modify pre-mRNA splicing (splice morpholino) in order to knockdown a gene function (Nasevicius & Ekker, 2000). The ATG morpholinos act by blocking the initiation of a protein translation at the ribosomes, thereby rendering the embryos devoid of a particular protein (Bill et al., 2009). Thus in order to investigate the function of myt1l on the development of the zebrafish, morpholino antisense oligonucleotides could be used to knockdown myt1l in vivo. Then, the outcome could be analysed by using in situ hybridisation, immunochemistry, and live imaging of possible morphological changes.

3. Identification of MYT1L target genes using a neuronal stem cell model

3.1. Introduction

Increasing amount of research has provided a better view on the possible function of *MYT1L*. We know, that in association with *Ascl1*, *Brn2* and *NeuroD1,* it transforms mouse and human fibroblasts directly into functional neurons (Z. P. Pang et al., 2011; Pfisterer et al., 2011; T. Vierbuchen et al., 2010). Combining this knowledge with the fact that Myt1l mRNA transcripts in rats were most abundant during prenatal brain development when most neurons are formed (J. G. Kim et al., 1997), it is hypothesised that it has a role in brain development. The complexity of the human brain has made it difficult to study many developmental genes in model organisms, highlighting the need for an *in vitro* model of human brain growth. In the present study we utilised human neural stem cells as a model to study potential gene targets of *MYT1L* and the impact of this transcription factor on the expression of other genes. In order to identify downstream targets of transcription factors, one of the key analyses is to identify gene expression changes, which occur when the function of a transcription factor of interest is perturbed. Research with lentiviruses enabling overexpression and/or short hairpin (sh) RNA knock-down provide a mean for interfering with the function of genes of interest. A virus is nature's own transport of genetic material, as it can incorporate into the cell without causing an immune response. Thus, viral infection is by far the most efficient transduction method (S. U. Kim, 2004). An important characteristic that distinguishes lentiviruses from the other viral vectors is their ability to transduce non-dividing cells, such as neurons, as well as dividing cells. Lentiviral vectors are able to stably integrate into the genome of host cells, without producing an immune response (Dissen et al., 2009). Further analyses to identify downstream target genes using a microarray gene expression analyses are described in this chapter.

3.1.1. Stem cells as an *in vitro* model for neurodevelopment.

Human embryonic stem cells have been widely used as model systems to inform scientists how typical developmental processes are implemented (Zeng et al., 2013). Their use in research settings proved to be very promising not only as candidate for future cell therapy but also for

studying normal neural development and entangling involvement of particular genetic factors during neurogenesis. Stem cells are defined based on two criteria: their ability to self-renew and ability to differentiate into different cell types. Stem cells can be further divided into different categories, according to their differentiation potential: totipotent stem cells, pluripotent stem cells, and multipotent stem cells. Totipotent stem cells are cells derived from the fertilised egg or morula stage embryo, and these cells are capable of differentiating into any type of cell (Mitalipov & Wolf, 2009). Pluripotent stem cells originate from the isolated inner cell mass of the developing blastocyst embryos (Evans & Kaufman, 1981). Pluripotent stem cells are able to differentiate into all three germ layers of the human body including ectodermal, endodermal and mesodermal cell types. Multipotent stem cells are cells isolated from specific tissues, from foetuses or adults, and these cells are able to generate differentiated cells of the same tissue origin (Mitalipov & Wolf, 2009).

Neural stem cells (NSCs) are self-renewing, multipotent cells that can give origin to neurons, astrocytes and oligodendrocytes and can be readily expanded *in vitro* (Mothe & Tator, 2012). During embryonic development, NSCs are found in several regions such as cerebellum, hippocampus, cerebral cortex, basal forebrain and spinal cord (Sally Temple, 2001). In the adult mammalian brain, it is generally accepted that NSCs are found in the dentate gyrus of the hippocampus and the subventricular zone (SVZ)(Doetsch, Caille, Lim, Garcia-Verdugo, & Alvarez-Buylla, 1999). This suggests the existence of a previously unrecognised neural plasticity in the mature CNS and the possibility of neural reconstruction in the adulthood. It has been suggested that the adult CNS retains the capacity for limited self-renewal in order to maintain its proper function, like learning and memory (Gage, 2000).

In vitro, NSCs can be maintained at their proliferative state by combining basic fibroblast growth factor (bFGF) and epidermal growth factor (EGF) in the defined or supplemented medium (Conti et al., 2005; Gage, 2000). Adherent NSCs provide an excellent platform to study genetic factors of neural development, as they can be easily expanded and readily manipulated. A method that has been widely used to maintain stable cultures of neural stem/progenitor cells is immortalisation. Cells are immortalised by using viral oncogenes, such as SV40 *largeT* or *v-myc* (S. U. Kim, 2004). These cell lines were shown to retain a lot of the characteristics of the NSCs, with the capacity to differentiate into neurons and glial cells allowing for an extensive study *in vitro* and also after transplantation. Nonetheless, the presence of a genetic alteration makes immortalised cell lines a controversial model, because of the modification in the gene expression that can lead to further mutations and production

of tumorigenic cells. However, recent reports suggested another system that could lead to normally developing neural cells. It employs *c*-mycER^{TAM} gene technology to immortalise embryonic stem cells (Littlewood, Hancock, Danielian, Parker, & Evan, 1995). The c-mycER^{TAM} transgene is better suited as an immortalising agent for clinical applications because c-Myc protein function is conditional on the presence of the tamoxifen metabolite, 4-hydroxytamoxifen (4-OHT) (Stevanato et al., 2009). To achieve conditional growth, a fusion protein is generated that consists of a growth promoting gene, *c*-myc, and a hormone receptor that is regulated by a synthetic drug, 4-hydroxy-tamoxifen (4-OHT). Pollock *et al.* (2006) have reported that in the absence of growth factors and 4-OHT, the cells differentiate into neurons and astrocytes. Recent findings suggest that c-mycER immortilised cells, derived from spinal cord, easily differentiate into V2 interneurons and motoneurons upon removal of growth factors and 4-OHT, thus inactivating the cMycER complex (Cocks et al., 2013). Those cells have been stably transplanted into rats without a tumorogenicity (Amemori et al., 2013; Cocks et al., 2013). Since the first stable hESC line was derived (Thomson et al., 1998), they have been considered as an excellent source of precursor cells to treat variety of conditions, such as motor neuron disease, spinal cord injuries, Alzheimer's disease and many more. However, the great attention has been given to the possible therapeutic use of stem cells. However, hESCs are equally important as a research tool for human developmental processes, both normal and diseased.

3.1.2. Use of lentiviruses to study gene function

Viruses are natural gene delivery systems, as they can target specific cells while avoiding immuno surveillance. Particular viruses have been selected as gene delivery vehicles because of their capacities to carry foreign genes and their ability to efficiently deliver these genes associated with an efficient gene expression. Viral vectors derived from retroviruses (including lentiviruses), adenovirus, adeno-associated virus, herpesvirus and poxvirus are mostly employed in current clinical gene therapy trials worldwide (Walther & Stein, 2000). By eliminating pathogenicity of a specific virus while retaining efficiency of the gene transfer, the system may be well suited for a successful gene delivery for clinical and non-clinical research. Among these vector systems, retrovirus vectors represent the most prominent delivery system, since these vectors have a high gene transfer efficiency and mediate high expression of the therapeutic gene (Robbins & Ghivizzani, 1998).

Lentiviruses are members of the large retrovirus family (Goff, 2001). Lentiviral vectors gained interest in the molecular biology field dues to their ability to infect many cell types, including non-dividing cells such as stem cells and neurons. They are composed of two copies of a linear single stranded RNA molecule of 7-12 kb, which are protected by a protein core inside the enveloped virus particle (Barquinero, Eixarch, & Perez-Melgosa, 2004). Three open reading frames (ORFs) are common to all retroviruses: gag (group associated antigen) encodes for structural proteins and forms the core, pol encodes for the proteins needed in a viral replication and integration, and env encoding for the envelope glycoproteins that mediate virus entry. All retroviral genomes are flanked by two long terminal repeat (LTR) sequences and contain *cis*-acting packaging sequences (phi, Ψ) that are needed for the correct incorporation of unspliced genomic RNA molecules into the new virus particles (Goff, 2001). The life cycle of retroviruses commences by binding to a receptor on the host cell via glycoprotein ligands of the virus. This is followed by fusion of the virus envelope and the membrane of the host cell which allows an entry. Immediately after an entry, reverse transcription takes place and a pre-integration complex is formed (PIC). Lentiviruses are unique due to the fact that they can infect undividing cells by entering the nucleus of a cell through the nuclear envelope and PIC (A. Pfeifer & Hofmann, 2009). Once the provirus enters the nuclear envelope, it integrates itself with a host genome. The most widely used vector system is derived from human immunodeficiency virus -1 (HIV-1) and a 3rd generation of this lentivector system has been developed for an optimised performance (Alexander Pfeifer, Lim, & Zimmermann, 2010). The general strategy applied during the production of viral particles that are replication-defective has been to eliminate all non-essential genes from the HIV-1 genome and separate *cis*- acting sequences from *trans*-acting particles (De Palma & Naldini, 2002). It is achieved by dividing the coding regions for structural and enzymatically active proteins (gag and pol-genes), the envelope construct and the retroviral genome into individual expression plasmids. In such systems, the structural proteins and the envelope construct are expressed from helper plasmids, where a gene of interest is located within the vector construct containing the *cis*- acting sequences required for the vector packaging, reverse transcription and integration (Dull et al., 1998). The commonly used 3rd generation lentiviral vector consists of four plasmids: a packaging construct containing gag and pol genes which are essential for production of lentiviral particles; an env - an envelope protein; a plasmid expressing Rev protein - that is necessary to achieve sufficient expression of unspliced vector genomic RNA; a transgene plasmid (vector transfer construct) and the promoter, as well as *cis*-acting sequences and regulatory elements. In addition, an important safety feature is present

in the vector, which is accomplished by a deletion of the promoter/enhancer sequences in the U3 region of the 3'LTR (generating self-inactivating (SIN) vectors). During reverse transcription the proviral 5'LTR is copied from the 3'LTR, therefore transferring the deletion to the 5'LTR, leading to transcriptional inactivation of the provirus (Dull et al., 1998). High-titer lentivector preparations are essential to achieve high transgenesis rates. Lentiviral vectors are commonly produced by co-transfecting adherent HEK 293T cells that act as a packaging cell lines, with calcium phosphate (Follenzi & Naldini, 2002). A large scale production of lentiviruses for clinical or molecular research requires a high titer vector production. Currently, the lentivarial systems generate 10^5 – 10^6 transducing units (TUs) per ml, which can be increased by centrifugation. Lentiviral transduction has been demonstrated to be the most effective method of a gene transfer (20-25%) with the >95% cell viability (Cao et al., 2010). Electroporation and lipofection demonstrated very low levels of transfection efficiency. Considering that human embryonic stem (hES) cells are notoriously difficult to transfect, lentiviral- derived vectors are efficient tools for the delivery and stable expression of transgenes in hES cells.

To summarise, use of lentiviral vectors is an appropriate choice to study gene function due to their unique ability to transduce dividing and non-dividing cells, such as stem cells, and to stably integrate into the host cell genome and achieve a long term transgene expression *in vitro* and *in vivo* (Cockrell & Kafri, 2007).

3.1.3. Molecular characterisation of MYT1L and its role in neurodevelopment

Information about the role of *MYT1L* in neurodevelopment comes from clinical studies, where aberrations of this transcription factor were found to be associated with several neurodevelopmental disorders, such as schizophrenia (Y. Lee et al., 2012; Van Den Bossche et al., 2013; Vrijenhoek et al., 2008), intellectual disability (Rio et al., 2013; S. J. Stevens et al., 2011), major depressive disorder (T. Wang et al., 2010) and autism (Matsunami et al., 2013; K. J. Meyer, M. S. Axelsen, V. C. Sheffield, S. R. Patil, & T. H. Wassink, 2012).

So far, our knowledge about MYT1L is limited. We know, that the MYT1L gene (also known as NZF1) maps to chromosome 2 in humans (2p25.3), it spans >542kb and has 25 exons. It has 13 known transcripts in humans (http://www.ensembl.org/Homo_sapiens/Gene/), 4 of which are protein coding variants, the longest consisting of 1186 amino acids. *MYT1L* is a member of the myelin transcription factor protein family, which has two other members: *MYT1* and *ST18* and

is highly homologous to *MYT1*. MYT1 is a representative of the Cys-Cys-His-Cys (CCHC) zinc-finger protein family, which has been highly conserved during evolution between both species and family members. It has been suggested that MYT1/MYT1L binding can result in the recruitment of histone deacetylases (HDACs) to target gene promoters, which results in a transcriptional repression (Romm et al., 2005). These studies suggest that MYT1L may have a regulatory role in the developing nervous system. The role of *MYT1L* in regulating neuronal differentiation is further highlighted by breakthrough studies showing that this protein in combination with other factors, such as *ascl1* and *brn2* (also known as *pou3f2*) could directly convert mouse fibroblasts into neurons (T. Vierbuchen et al., 2010) Another study, found that addition of *POU3F2* to *MYT1L* and *ASCL1* can generate functional neurons from human pluripotent stem cells as early as 6 days after the transgene activation (Z. P. Pang et al., 2011). When combined with *NeuroD1*, these factors could also convert foetal and postnatal human fibroblasts into induced neuronal cells showing typical neuronal morphologies and expressing multiple neuronal markers. These induced neuronal cells express multiple neuron-specific proteins, generate action potentials, and form functional synapses (Z. P. Pang et al., 2011; T. Vierbuchen et al., 2010). It was reported that miR-9* and miR-124 were found to induce the conversion of human fibroblasts into neurons with *NeuroD2* (Yoo et al., 2011). Addition of *MYT1L* and *ASCL1* further enhanced the rate of conversion and the maturation of the converted neurons. All of the aforementioned studies have provided evidence to suggest the importance of *MYT1L* in neural development.

Another important finding comes from a collaboration between Dr Sylvane Desrivières (supervisor) and Professor Jack Price's laboratory, where they have tested, using quantitative real-time PCR (qPCR), the expression patterns of *MYT1L* in differentiated and undifferentiated human stem/progenitor cell lines (unpublished)(Figure 3.1).

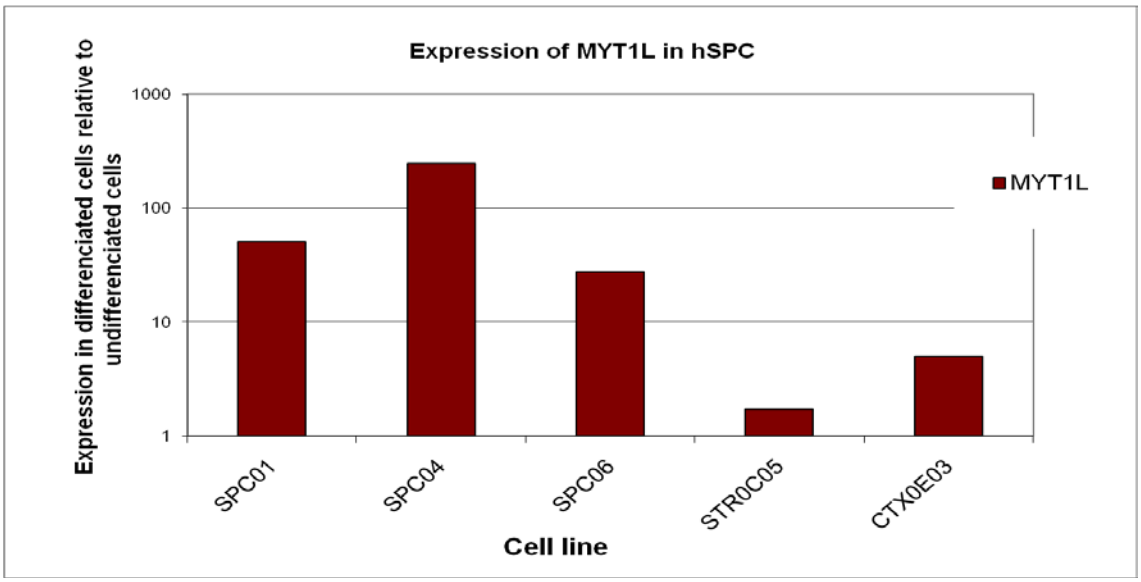


Figure 3.1. *MYT1L* expression in various human stem/progenitor cell lines. The y-axis shows the fold MYT1L mRNA induction in differentiated cells compared undifferentiated cells; the x-axis represents different types of human stem/progenitor cells. SPC stands for spinal cord embryonic stem cells, CTX stands for embryonic cortical stem cells, and STR stands for cells derived from human foetal striatum. Numbers in each cell line represent different clones.

These experiments indicated that *MYT1L* has very low or no expression in undifferentiated stem cells and is strongly induced during the differentiation, in most cell lines tested. This suggests that *MYT1L* may play a role in the differentiation process. As the SPC-04 cell line showed the highest induction of *MYT1L* upon differentiation (>200 fold), we planned to use this cell line for subsequent neural differentiation studies (described below).

3.2. Aims.

Despite its apparent involvement in neurogenesis, the cellular and molecular functions of *MYT1* remain largely unknown. The role of *MYT1L* on the expression of other genes during neural development is yet to be determined. The search for interacting partners of MYT1L will help us to understand the biological role of MYT1L and mechanisms by which its gene product regulates neural differentiation. The general aim of this chapter is to functionally characterise *MYT1L* and investigate its direct gene targets using different approaches to achieve two main points:

1. Generate a *MYT1L* over-expressing lentiviral construct for infection into undifferentiated SPC-04 cell line to enable co-immunoprecipitation of exogenous MYT1L protein followed by mass spectrometry analyses to identify MYT1L-interacting partners.

2. Use a MYT1L knock-down lentiviral construct to infect pre-differentiated neuronal stem cells to enable:
    a. Identification of downstream transcriptional targets of *MYT1L* by comparing global gene expression levels between *MYT1L* knock-down cells and control cells differentiated for 7 and 14 days using microarray technology.
    b. Characterisation of differentially expressed genes by performing a gene ontology term enrichment analyses and Ingenuity Pathway Analysis to identify possible biological processes and molecular mechanisms by which *MYT1L* might influence brain development.

3.3. Materials and methods.

3.3.1. Construction of lentiviral vectors for overexpression of MYT1L

This section describes the constructs used to develop overexpression plasmids. Several commercially available systems were used in this attempt; however none of them produced satisfactory results meaning we were unable to successfully express exogenous *MYT1L*. Results and possible implications will be discussed later in the chapter.

Two lentiviral vectors were constructed, one to express MYT1L in its native form and a second one as a recombinant protein with a C-terminal V5 epitope, allowing recognition of the recombinant protein by an anti-V5 antibody. pLenti7.3/V5-TOPO® vector (Life Technologies, UK) was used for cloning MYT1L to produce a recombinant fusion protein. pLenti7.3/V5-TOPO vector is a lentiviral expression vector that has been adapted for use with TOPO TA Cloning technology. TOPO cloning is a method for cloning PCR products into a plasmid vector. The key to TOPO cloning is the enzyme DNA topoisomerase I, which functions both as a restriction enzyme and as a ligase. The plasmid (pLenti7.3/V5-TOPO) is provided as linearised vector with a single 3′ thymidine (T) overhangs for the TOPO Cloning and the topoisomerase I covalently bound to each 3′ phosphate. The 3'A overhangs of the PCR product complement the 3'T overhangs of the vector and allow for fast ligation with the topoisomerase I. Using the TOPO cloning technology, the pLenti-TOPO vectors are a highly efficient cloning strategy for the direct insertion of amplified PCR products into a plasmid vector. Once cloned, the plasmid can then be transformed into competent bacterial cells. After analysis of the plasmid, and transfection, the PCR product can be directly expressed in chosen cell lines. Its schematic is

shown on Figure 3.2, illustrating the main features of the vector including the human cytomegalovirus (CMV) immediate early promoter to control high-level expression of the gene of interest, the WPRE (Woodchuck Posttranscriptional Regulatory Element) from the woodchuck hepatitis virus, which permits increased transgene expression, the Emerald Green Fluorescent Protein (EmGFP) to monitor transduced cells and the V5 epitope which allows detection of the recombinant fusion protein by anti-V5 antibodies.

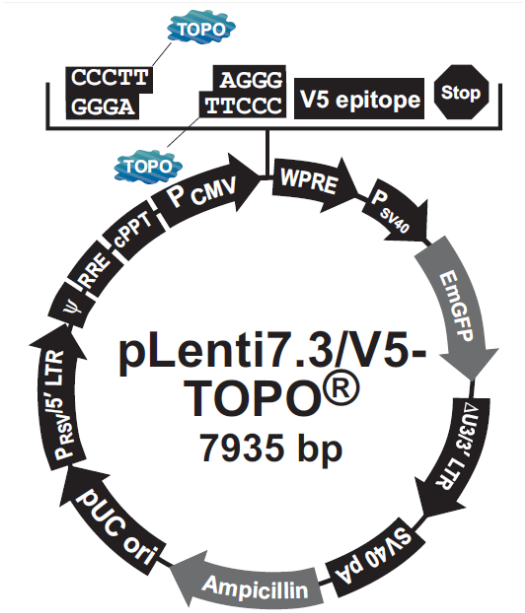


Figure 3.2. Schematic representation of pLenti7.3/V5-TOPO® vector. This lentiviral expression plasmid contains a CMV promoter, an EmGFP reporter gene, an ampicillin resistance gene. A V5 epitope is present in the multiple cloning site which can result in a tagged protein of interest if cloned in frame with the V5 sequence.

3.3.1.1. Primer design for MYT1L amplification

Primers were designed to amplify MYT1L ORF (open reading frame) by PCR using the Phusion® High-Fidelity DNA Polymerase (Thermo Scientific). The forward primer incorporated the KOZAK consensus ([G/A]NNATGG**)** sequence – the optimal sequence required for translation initiation. Two reverse primers were designed to obtain either the MYT1L protein sequence in frame with the V5 epitope tag resulting in a fusion protein (MYT1L +V5) or the native MYT1L protein. Table 3.1 shows primer sequences used for the amplification. To obtain the native form of MYT1L, a stop codon (TGA) was included in the reverse primer. Whereas to obtain a V5-tagged MYT1L protein, a stop codon was removed and additional (underlined) nucleotides were included to clone in frame with the V5 epitope.

Table 3.1. Primers used to amplify MYT1L ORF for cloning into pLenti7.3/V5 TOPO® TA vector. In order to generate V5-tagged and native form of the protein, two different reverse primers were used. Both primers have a similar sequence except that the primer resulting in the fusion protein has no stop codon (highlighted in red in the native primer) and has two additional base pairs (guanine and cytosine – underlined) to keep the MYT1L sequence in frame with the V5 sequence. Kozak consensus sequence is underlined in the forward primer.

Primer	Sequence
Forward	5' – AAG**ATG**GAGGTGGACACCGAGGA – 3'
Reverse with V5 tag (MYT1L+V5)	5' – GCGACCTGAATTCCTCTCACAGC – 3'
Reverse native	5' – TCAGACCTGAATTCCTCTCACAGC – 3'

3.3.1.2. PCR amplification of MYT1L

The PCR reaction to amplify *MYT1L* ORF was performed using the Phusion® High-Fidelity DNA Polymerase (Thermo Scientific, UK), in a total volume of 50µl containing 5x Phusion high fidelity (HF) buffer (Thermo Scientific, UK), 10mM dNTP mix, 0.5µM forward and reverse primer (either native or V5), 1U Phusion DNA Polymerase, 10ng DNA template and dH_20 to a final volume of 50µl. The thermal cycler conditions are presented in Table 3.2:

Table 3.2: PCR thermal cycling conditions to amplify MYT1L ORF.

Step	Time	Temperature (°C)
1 - Initial denaturation	10s	98
2 - denaturation	10s	98
3 - annealing	30s	68
4 - extension	2mins	72
5 - number of cycles	Go to step 2 - 35 times	
6 - final extension	10mins	72

3.3.1.3. DNA gel electrophoresis

PCR products were electrophoresed, in order to evaluate the success and specificity of the amplification; i.e., to confirm the presence of a single, clear DNA band representing the *MYT1L* ORF (~3.55kb in length). DNA gel electrophoresis was performed using a 0.8% agarose gel, which consisted of 0.8g agarose and 1x Tris/Borate/EDTA (TBE) buffer to a total volume of 100ml in a conical flask. The mixture was swirled gently by hand and heated in the microwave (on the high setting) until the agarose dissolved completely. The solution was then placed under a cold tap for ~1min. Once cooled, 2µl ethidium bromide was added and the solution was mixed thoroughly by gentle swirling. Ethidium bromide is a fluorescent intercalating dye that can be seen under UV light and is used to visualise DNA on agarose gels. The solution was then poured into a sealed (with an autoclave tape) gel casting tray containing a comb to form appropriately sized wells and left to set for ~20mins. Wells were loaded with samples consisting of 5µl DNA and 1µl 6x gel loading buffer (GLB) (Thermo Scientific, UK), with one or two wells reserved for 5µl of 1kb DNA ladder (Invitrogen). Gels were electrophoresed at 110V for ~1hr or until the ladder had undergone sufficient separation. Gels were then visualised using a UV transilluminator.

3.3.1.4. Preparation of lysogeny broth (LB) and LB agar

LB was prepared by dissolving 20g Lennox L Broth Base (Invitrogen) in dH_2O to a total volume of 1 litre, whereas LB agar was prepared using 32g Lennox L Agar (Invitrogen) with dH_2O to a total volume of 1 litre. Both solutions were autoclaved. The pLenti 7.3/V5 vector expresses resistance to the antibiotic ampicillin (Figure 3.2), therefore 100µg/ml ampicillin was added to cooled, autoclaved LB agar and LB broth. This procedure ensures that only bacteria containing successfully ligated plasmids would be present, as open plasmids are unable to express ampicillin resistance. Directly after the addition of ampicillin, the LB agar solution was distributed evenly into numerous Petri dishes (around 10ml per dish) and plates were left to cool at room temperature, and then stored at 4°C. LB agar plates were sealed to prevent contamination.

3.3.1.5. Ligation of MYT1L into pLenti 7.3/V5 TOPO expression vector

TOPO Cloning technology allows a quick production of PCR product, by ligating it into the pLenti-TOPO vector, and followed by a transformation of the recombinant vector into stable, component cells. pLenti 7.3/V5 TOPO® TA cloning kit (Life Technologies, UK) was used to ligate a fresh PCR product with the expression vector.

The TOPO® cloning reaction consisted of: 4µl of the PCR product, 1µl salt solution, 1µl pLenti-TOPO® vector in a final volume of 6µl. In parallel, a control reaction was set up with no insert DNA to assess the re-ligation efficiency of the linearised vector. This mixture was incubated at RT for 20 minutes and then placed on ice. 5µl of this reaction was then transformed into MAX Efficiency Stbl2™ Competent *E.coli*, according to manufacturer's instructions.

3.3.1.6. Transformation of bacteria with plasmid DNA

5 µl of ligated products (V5-MYT1L, Native MYT1L and empty vector) were added to 50µl chemically competent MAX Efficiency Stbl2 cells (Invitrogen, UK) and incubated on ice for 30 mins. Bacteria were then inactivated by heat-shock at 42°C for 25s and returned to ice for 2 s. Next, 450 µl of S.O.C medium was added to each transformation and the bacteria were incubated at 30^{o}, shaking at 225rpm for 1.5 hours. The cells were then gently mixed by pipetting and 100µl of the cell suspension from each construct was streaked onto LB agar plates (containing 100µg/ml ampicillin), which were then incubated overnight at at 30°C to isolate individual colonies.

3.3.1.7. Screening bacterial colonies for positive clones

Individual colonies were aseptically picked in order to screen for positive clones (vector with insert) and were pre-cultured in 5 ml of LB Broth media (Invitrogen, UK) containing 100µg/ml ampicillin at 30^{o} C overnight and used for small- or large-scale plasmid production.

3.3.1.8. Plasmid purification

For small-scale plasmid isolation minipreps were used (NucleoSpin Plasmid Kit, Macherey-Nagel, Germany). The tubes containing the 5 ml of pre-culture bacteria were centrifuged at

1500 RPM for 5 minutes at room temperature. Approximately, 4 ml of the supernatant was discarded using sterile tips, and bacteria were resuspended in the remaining 1ml and transferred to a microcentrifuge tube for further centrifugation at 11000 x *g* for 1 minute. The supernatant was discarded and the bacterial pellet was resuspended in 250 µl of buffer A1 (containing RNase A at concentration of 100 µg/ml). This was followed by the addition of 250 µl of lysis buffer A2, the suspension was mixed by inversion (6 to 8 times), incubated at room temperature for 5 minutes and the lysis was neutralised by the addition of 300 µl of buffer A3. The mixture was centrifuged at 11000 x *g* for 5 minutes to pellet bacterial debris and the plasmid-containing supernatant was carefully removed and transferred into a NucleoSpin Plasmid column provided with the kit. Plasmid DNA was bound to the column by brief centrifugation (1 min) at 11000 x *g.* Column-bound DNA was washed with 600 µl of buffer A4 (containing 80% ethanol), flow-through discarded and the membrane was dried by centrifugation for 2 mins at 11000 x *g.* Plasmid DNAs were eluted with 30 µl of buffer AE (5 mM TrisHCl,pH 8.5).

For a large-scale plasmid preparation, 1 ml of the pre-culture bacteria was added to 400 ml of LB Broth media (supplemented with 100 µg/ml ampicillin) and incubated at 30^{o} C overnight while shaking at 50RPM. The bacteria were then harvested by centrifugation at room temperature for 20 mins at 1500 RPM. Supernatants were discarded and the bacterial pellets used for maxi-prep DNA isolation using the Endofree Plasmid Maxiprep Kit (Qiagen, UK), according to the manufacturer's instructions. Briefly, bacteria were resuspended in 10 ml buffer P1 [50 mM TrisHCl (pH 8.0) and 10 mM EDTA] supplemented with RNase A (100 µg/ml). The mix was lysed by adding 10 ml of buffer P2 (200 mM NaOH and 1% SDS) and incubating for 5 mins at room-temperature. Next, lysates were neutralised by the addition of 5 ml of chilled buffer P3 (3.0 M potassium acetate,pH 5.5) and mixed by a vigorous tube inversion (6-8 times). Then, the lysates were filtered through the QIAfilter cartridge, 2.5 ml of buffer ER was added for endotoxin removal and lysates were incubated on ice for 30 mins. The supernatant was then applied into an equilibrated QIAGEN-tip 500 column and allowed to flow-through by gravity flow. The column was then washed twice with 30 ml of buffer QC (1 M NaCl, 50 mM MOPS (pH 7.0) and 15% isopropanol). The plasmid DNAs were eluted in 15 ml of buffer QN (1.6 M NaCl, 50 mM MOPS, pH 7.0 and 15% isopropanol), precipitated by adding 10.5 ml of isopropanol (Sigma, USA), which was followed by centrifugation at 20000 x *g* for 30 min at 4^{o}C. The supernatants were carefully decanted and DNA pellets were washed with 5 ml endotoxin-free 70% ethanol and centrifuged further for 10 min at 20000 x *g* at 4^{o}C; Again, supernatants were carefully removed, without disturbing the pellet, and the plasmid DNA pellets were air-

dried for 5 minutes and resuspended in 150 µl of TE buffer (10 mM Tris-HCl,pH 8.0 and 1 mM EDTA). DNA concentrations were measured by UV spectrophotometry using a nanospectrophotometer (NanoDrop 1000, Thermo Scientific) and plasmids were stored at -20^{o} C till required.

3.3.1.9. Verification of correct DNA inserts

Following transformation, colony isolation and plasmid purification by minipreps, samples were screened for positive clones by restriction digest using the HindIII restriction enzyme. This enzyme was chosen for verification, as it cuts 6 times in the pLenti vector (Figure 3.3) and it cuts the MYT1L ORF once creating two fragments: 3118bp and 437bp. HindIII cuts are not within the multiple cloning site of the vector, they are placed before the CMV promoter site and GFP site, making it a good candidate for testing the orientation of the insert. The TOPO cloning site is located at 2558 – 2567bp, and the V5 epitope site is located at 2630-2671bp on the vector. If the vector had no insert present, the expected sizes after cutting with HindIII would be as follows: 312 bp, 481bp, 556bp, 584 bp, 2658bp and 3344bp. If *MYT1L* was inserted in a correct (sense) orientation (5' towards 3') instead of a 2658bp band we expected two bands sized: 2353bp and 3747bp representing the *MYT1L* insert. The restriction reaction was performed as follows: 10 µl of plasmid DNA from the mini-prep was digested with 5U of HindIII enzyme (NEB, UK) in the presence 1x NEB restriction enzyme buffer 2 (NEB, UK) and topped up with dH_20 to a final volume of 20 µl. The reaction was inclubated for 1.5h at 37 ^{o}C. The DNA digestions were separated by electrophoresis on a 1.5% agarose gel (Sigma, UK) made up with 1X TBE buffer (0.0089 M Tris, 0.089 M Borate, 0.002 M EDTA; Merck, Germany) and supplemented with 2% of Ethidium Bromide (Electron, UK). 10 µl of the digested samples were mixed with 2 µl of 6X gel-loading buffer (Thermo Scientific, UK) and loaded into the gel wells. Additionally, two wells (first and last well on the gel) were used for loading the 1 kb DNA ladder (250-10,000 bp Sigma, UK). Electrophoresis was at 100 V for 1.5h. The DNA bands obtained from the gel were visualised using ultraviolet light. In addition, DNA plasmids were sent for sequencing (Bioscience, UK) to check the correct *MYT1L* ORF and to verify that no change of the sequence occurred during cloning process.

General Description

DNA pLenti7.3/V5TOPO_verA

Entire molecule length: 7935 bp

Restriction/Methylation Map

Enzyme	# of cuts	Positions
ApaLI	3	2905 5966 7212
AvaI	3	1700 2603 3621
BamHI	1	2529
ClaI	1	1924
HindIII	6	308 864 1448 1929 4587 7931
NcoI	4	2244 3104 3507 4964
PstI	1	2579
SmaI	1	3623
XmaI	1	3621

Figure 3.3. Figure represents restriction sites for pLenti 7.3/V5-TOPO vector. TOPO cloning site within pLenti vector is located at 2558-2567bp HindIII enzyme cuts vector six times and was used to determine orientation of the inserted MYT1L sequence.

3.3.1.10. Transfection and virus production

To produce a lentiviral stock, 293FT cells (Invitrogen, UK) were used. Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM), containing 10% heat-inactivated Foetal Bovine Serum (FBS), 4mM L-Glutamine, 1mM MEM Sodium Pyruvate, 0.1 mM MEM Non-Essential Amino Acids and 1% penicillin/streptomycin and 500 µg/ml geneticin (Sigma, UK). Cells were transfected using ViraPower™ HiPerform™ Lentiviral TOPO Expression Kits (Invitrogen, UK) and Lipofectamine™ 2000 (Life Technologies, UK) according to the manufacturer's instruction.

Briefly, the day prior transfection, 293FT cells were seeded at a density of 5.5 x 10^6 cells/ml in 10cm sterile dishes in 10ml of culture medium without antibiotics. On the day of transfection, the culture medium was replaced with 5ml Opti-MEM I medium containing 5% FBS. Three separate reactions were prepared: empty pLenti vector, pLenti + *MYT1L* native and pLenti + V5-tagged *MYT1L.* For each transfection sample, 9µg of the ViraPower™ Packaging Mix (Life Technologies, UK) and 3µg of appropriate pLenti expression plasmid DNA (empty vector, MYT1L native and MYT1L-V5) was added to 1.5ml Opti-MEM I Medium without serum and mixed gently. Separately, for each transfection reaction, 36µl of Lipofectamine™ 2000 was diluted in 1.5ml Opti-MEM without serum, mixed gently and incubated at RT for 5 minutes. The diluted Lipofectamine™ 2000 was subsequently added to the diluted plasmid DNA mix. This combination was then incubated for 20 minutes at RT. Following incubation, all the DNA-Lipofectamine™ 2000 complexes were added dropwise to the plate of 293FT cells, mixed

gently and incubated at 37°C with 5% CO_2. The following day, the medium containing the DNA-Lipofectamine™ 2000 complexes was removed and replaced with 10ml culture medium without antibiotics and placed back in an incubator at 37°C with 5% CO_2 for a further 24-48h. The plates were microscopically examined for the presence of GFP (fluorescent reporter), an indicator of transfection efficiency. Supernatants containing virus were collected 24 and 48h post transfection. These were centrifuged at 3000RPM for 15 minutes at RT to pellet the cell debris. Following centrifugation, the supernatants were filtered through 0.22µm pore nitrocellulose filter (Millipore, UK), aliquoted and stored at - 80°C until use.

3.3.1.11. Viral transduction

293T and 293 FT cells were infected with viruses overexpressing MYT1L as well as the empty vector as a control. The day before transduction, 5 x 10^5 293T cells were plated in each well of a 6-well plate and incubated overnight at 37°C with 5% CO_2. On the following day, the media was removed and changed to 1 ml of serum free and antibiotic free cell media and 2 ml of lentiviral supernatant with a final concentration of 0.4mg/ml Polybrene (Sigma, USA) to increase transduction efficiency. The plates were centrifuged for 90 mins at 2500 RPM and incubated for 2.5h at 37^o C in 5% CO_2. The media was then changed to 3 ml of complete media and the incubation continued for a further 72 hours. Cells were observed daily under a fluorescent microscope for the presence of GFP and estimation of transduction efficiency. Additionally, to accurately assess transduction efficiency, cells were lysed and RNA was extracted to perform qPCR.

3.3.1.12. Total RNA extraction.

All RNA extractions were performed using the RNeasy Mini Kit (Qiagen, UK), according to the manufacturer's instructions. Briefly, the culture medium was aspirated and cells were washed twice with cold 1x PBS, then 600µl of RLT buffer containing 1% β-mercaptoethanol (Sigma, UK) was added to each plate. Cells were then collected using a sterile cell scraper and transferred to a 1.5ml tube. The mixture was vortexed briefly to ensure thorough lysis. All cell lysates were then homogenised using the QIAshredder columns (Qiagen, UK), and centrifuged at 21000 x *g* for 2 minutes. One volume of 70% molecular biology-grade ethanol (Sigma, UK) was added to the lysates and mixed well by pipetting. The mixture was then transferred into an RNeasy spin

column and centrifuged for 15s at 21000 x *g*. The flow-through was discarded and 350µl of RW1 buffer was added to each column followed by centrifugation for 15s at 21000 x *g*. Again, the flow was discarded and 80µl of DNase I incubation mix, (1:7 ratio of DNase I stock solution and Buffer RDD; Qiagen, UK) was added to each column. The reaction was incubated for 15 mins at room temperature, followed by addition of 350µl of RW1 buffer. The columns were then centrifuged for 15s at 21000 x *g.* The flow was discarded and each column was washed with 500µl of RPE buffer followed by centrifugation for 15s at 21000 x *g*, discarding the flow at the end. A second wash with 500 µl of the RPE buffer was applied and the column was centrifuged for 2mins at 21000 x *g*. Next, the column was transferred into a clean RNase-free collection tube and RNA was eluted by adding 30µl of RNase-free water directly to the membrane of the column and centrifuging it for 1 minute at 21000 x *g*. RNA samples were quantified using a nanospectrophotometer (NanoDrop 1000).

3.3.1.13. Reverse transcription and complementary DNA (cDNA) synthesis

Reverse transcription was performed using the Invitrogen SuperScript™ III first-strand cDNA synthesis kit (Invitrogen, UK). 1µg of total RNA was added to 1 µl of 10mM dNTP mix and 1µl of Oligo$(dT)_{18\text{-}20}$ (0.5µg/µl) and the volume made up to 10µl with RNAse-free water. Each sample was incubated at 65°C for 5 minutes and then placed directly on ice. Next, to each sample 2µl of 10x RT buffer (200 mM Tris-HCl,pH 8.4, 500 mM KCl), 4µl of 25mM $MgCl_2$, 2µl of 0.1M DTT, 1µl of RNaseOUT™ Recombinant RNase Inhibitor and 1µl of Superscript™ III RT (2000U/µl) was added. The samples were incubated at 50°C for 50 minutes and the reaction was terminated by incubation for 5 minutes at 85°C. 1µl of RNase H was added to each sample and then samples were incubated at 37°C for 30 minutes. cDNA was then diluted with distilled water, making 100µl of each sample at a concentration of 10ng/µl.

3.3.1.14. Real Time quantitative PCR.

cDNA samples were amplified using an ABI Prism 7900HT sequence detection system in a final volume of 10µl containing 2x power SYBR green master mix (Applied Biosystems, UK), 4µl of cDNA (10ng/µl) and 0.07µM of forward and reverse primer (Table 3.3). PCR reactions were performed in triplicate under the following thermal cycler conditions: 95°C for 15 minutes followed by 95°C for 30 sec and 59°C for 30 sec for 40 cycles. The PCR reaction was evaluated

by dissociation curve analysis to ensure that the PCR product generated were specific. *β-ACTIN* was used as the housekeeping gene against which expression values for *MYT1L* were normalised to produce ΔCt values ($\Delta Ct = Ct_{myt1l} - Ct_{actin}$).

Table 3.3. Primer sequences used for qPCR amplification of *MYT1L* and the control gene *β-ACTIN* in 293T and 293FT cells.

Primer	Sequence
h β-ACTIN Forward	5' – GCTCGTCGTCGACAACGGCTC – 3'
h β-ACTIN Reverse	5' – CTCTGGCACGCTCGAACT – 3'
hMYT1L Forward	5' – TGGAGAGCAACCTGAAGACC – 3'
hMYT1L Reverse	5' – CAAACATGATCTGGGTCATCTTCTC – 3'

3.3.2. Lentiviral-mediated gene silencing of MYT1L

3.3.2.1. Lentiviral constructs for MYT1L silencing

Concentrated lentiviral stocks carrying either non-silencing shRNA or *MYT1L* silencing shRNA were kindly provided by Lourdes Martinez-Medina (LMM) (SGDP, King's College London). The *MYT1L* shRNA was cloned into a lentiviral vector that belonged to the Thermo Scientific Open Biosystems pGIPZ shRNAmir library developed by Dr Greg Hannon and Dr Steve Elledge (Figure 3.4). This particular gene silencing system combines the advantages of micro-RNA adapted shRNA with the pGIPZ lentiviral vector to create a highly active RNAi pathway with its unique short hairpin design. The short hairpin is expressed as human microRNA-30 (miR30) primary transcript. The hairpin stem includes 22 nucleotides of dsRNA and the miR30 loop composed of 19 nucleotides. Additionally, the miR30 design allows the use of rule based design such as destabilising the 5' end of the antisense strand to incorporate strand specific microRNAs into the RISC machinery.

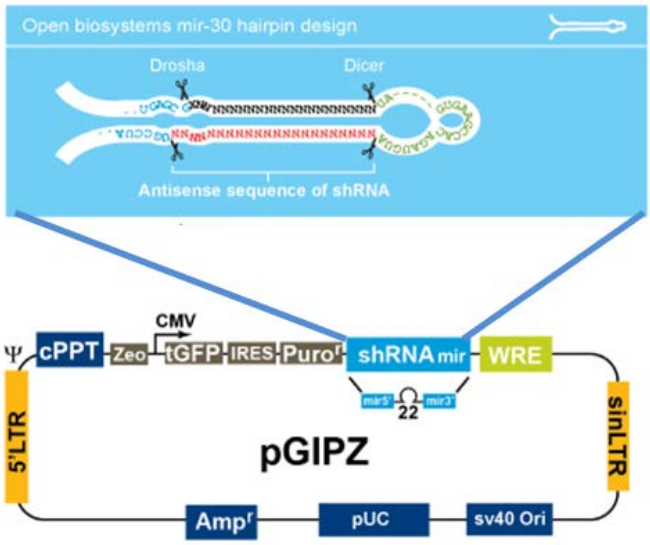


Figure 3.4. A schematic representation of pGIPZ vector (bottom) and a schematic shRNA design (top) that is cloned inside of the vector. The pGIPZ lenti-vector contains various features to allow effective expression and tracking of shRNAs, such as bacterial selection (ampicillin (Amp^r), zeocin (Zeo) or puromacin (Puro^r resistance); CMV promoter, tGFP marker to track shRNA expression and the shRNAmir. The short hairpin RNA constructs are composed of 22 nucleotides (red) with a complementary strand (black) that forms a dsRNA and a miR-30 loop (green). It was designed to have a Drosha and a Dice processing site to achieve greater siRNA production (figures from Thermo Scientific Open Biosystems pGIPZ shRNAmir library, UK)

The sense sequences for the shRNAs used are provided below:

1) pGIPZ + non-silencing shRNA; sense sequence: 5' – ATCTCGCTTGGGCGAGAGTAAG – 3'

2) pGIPZ + shRNA *MYT1L* 1; *MYT1L* target sequence: 5' – CCGTGACTACTTTGACGGAAAT – 3'

Open Biosystems TransLenti Viral Packaging System (Thermo Scientific, UK) was used for transfection and production of lentiviral particles according to the manufacturer's instructions. The lentiviral supernatants were filtered through a sterile 0.22µm pore filter (Sartorius Stedim, UK) and concentrated by ultracentrifugation at 75000 *x g* (Beckman Coulter, UK; SW27 Rotor) at 4°C for 1.5 hours. The supernatant was carefully removed without disturbing the viral pellet, and then cold PBS (30µl) was added to the vial and stored overnight at 4°C. The following day a pellet was resuspended, aliquoted and stored at -80°C.

3.3.2.2. SPC-04 culture and maintenance

SPC-04 cells were maintained in reduced modified media (RMM-) with growth factors and 4-hydroxy-tamoxifen (4-OHT) denoted (RMM+) at 37°C in a humidified chamber at 5% CO_2. Components and final concentrations required for RMM- and RMM+ preparation are detailed in Table 3.4:

Table 3.4. RMM media preparation for SPC-04 cell line. RMM+ includes all the components with groth factors (highlighted in grey), where RMM- includes only non-highlighted components

Components for RMM	Final concentration	Stock preparation
DMEM:F12 (Ham) (Life Technologies, UK)	-	no preparation required
Albumin serum, Human (HAS) (Baxter Healthcare Ltd., UK)	0.03%	stock: 20% (w/v), no preparation required
Human Apo-Transferrin (Sigma, UK)	100µg/ml	50mg/ml diluted in DMEM:F12with 1mM HEPES pH 7.4
Insulin, Human recombinant (Sigma, UK)	5µg/ml	stock: 10mg/ml, no preparation required
L-Glutamine (Sigma, UK)	2mM	stock: 200mM, no preparation required
Progesterone (Sigma, UK)	60ng/ml	1st step: 1mg (stock) dissolved in 100µl molecular biology grade ethanol 2nd step: dilute the mix in DMEM:F12 to final concentration 20mg/ml
Putrescine Dihydrochloride, Human (Sigma, UK)	16.2µg/ml	stock: 8.1mg/ml dissolved in tissue culture grade water
Sodium Selenite (Sigma, UK)	40ng/ml	stock: 20µg/ml dissolved in DMEM:F12
4-hydroxy-tamoxifen (Sigma, UK)	100nM	stock: 1mM dissolved in molecular biology grade ethanol
Human EGF (PeproTech EC Ltd, Germany)	20ng/ml	stock: 10µg/ml dissolved in DMEM:F12 with 1mM HEPES and 0.03% human albumin serum
Human bFGF (PeproTech EC Ltd, Germany)	10ng/ml	stock: 10µg/ml dissolved in DMEM:F12 with 1mM HEPES and 0.03% human albumin serum

Reduced modified media plus (RMM+) was composed of DMEM:F12 (Gibco, UK) supplemented with 4-OHT (100nM; Sigma, UK) and all the components written in Table 3.4. The media was filtered through a 0.2µM Stericup filter (Millipore, UK) and stored at 4^{o}C for a maximum of 4

weeks. Differentiation media RMM- was prepared in the same manner but this media was depleted of the growth factors and 4-OHT.

3.3.2.2.1.Preparation of differentiation inducing agents

Two reagents were used in combination with RMM- medium to induce differentiation:

1) DAPT (LY-374973 N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester) at stock concentration 10mM in DMSO.

2) ATRA (All trans retinoic acid) stock concentration 10mM in DMSO.

Both reagents were aliquoted and stored at -80°C.

DAPT is an inhibitor of the γ-secretase complex (Dovey et al., 2001). ITacts as an inhibitor of Notch (Sastre et al., 2001), a γ-secretase substrate. it has been found to promote neuronal differentiation from human pluripotent stem cells (Androutsellis-Theotokis et al., 2006; Chambers et al., 2012; Dimos et al., 2008). Another study demonstrated that DAPT-mediated inhibition of the Notch response resulted in an enhanced neuronal differentiation (Crawford & Roelink, 2007). ATRA is the oxidised form of Vitamin A and is a signalling molecule involved in pathways that control differentiation and proliferation (Duester, 2008; Ertesvag, Naderi, & Blomhoff, 2009). ATRA has been widely used for various differentiation protocols including neural differentiation from human embryonic stem cells (Dhara & Stice, 2008; Sasai, 2002; Takahashi, Palmer, & Gage, 1999; Wichterle, Lieberam, Porter, & Jessell, 2002).

3.3.2.2.2.Maintenance and passaging of SPC-04 cells

All flasks and plates used for SPC-04 cell culture and experiments were Nunclon™ Δ Delta (Nunc) surface treated to ensure optimal conditions for cell attachment and growth. Cells were grown until 80% confluency. All reagents were pre-warmed to 37 °C.

Firstly, tissue culture flasks were coated with DMEM containing 20µg/ml mouse laminin, for at least 3h. Just prior to plating cells, the excess laminin was removed and the flask was washed twice with DMEM:F12 to remove residual laminin. DMEM:F12 was replaced with RMM+ and the flask was returned to the incubator. Next, culture flask to be passaged was first rinsed with HBSS (Hank's Balanced Salt Solution without Ca2+ and Mg2+; Life Technologies), followed by

immediate addition of TrypZean (Lonza, UK) and returned to the incubator for approximately 2 minutes till the cells started detaching. Then Trit (1% Human Albumin Serum, 0.55% soyabean trypsin inhibitor, and 0.025U/ml bensonase solution (Merck, UK) in DMEM:F12) was added to the flask to inactive the enzyme. The cell suspension was transferred using a sterile pipette to a 15ml Falcon tube, followed by centrifugation at 900RPM for 5 minutes. The pellet was then resuspended in 1-2ml of RMM+ medium, the number of cells counted and the cell suspension was used for further plating or freezing. For freezing, a 20% DMSO solution (diluted in RMM+) was prepared and 0.5ml of it was added to 0.5ml of cell suspension to give a final concentration of 10% DMSO. Cells at a final concentration of at least 2 x 10^6 cells/ml were frozen at 1°C/min in a freezing container (Nalgene) filled with isopropanol and stored at -80°C overnight and transferred to a liquid nitrogen the day after. Cells for further propagation were plated in freshly coated and prepared flasks at a density of 20000cells/cm^2.

3.3.2.2.3.Cell counting

After trypsinisation, cells were counted and their viability was assessed using NucleoCounter® NC-100 (Chemometec, Denmark). For total cell number, an equal amount of cell suspension, Reagent A – lysis buffer (Chemometec, Denmark) and Reagent B - stabilising buffer (Chemometec, Denmark) were mixed together in a final volume of at least 120µl. The mixture was loaded into a NucleoCassette™ and placed into the NucleoCounter® to obtain the cell concentration. NucleoCassette contains propidium iodine (PI), an integrated fluorescent dye, which binds to the DNA released from the trypsinised cells. The fluorescence is then detected by the camera in the NucleoCounter and correlated into a cell count. Presented concentration (cells/ml) was multiplied by the dilution factor to obtain total number of cells.

3.3.2.2.4.Differentiation of SPC-04 cells

To induce differentiation, cells were seeded in a 6-well laminin- coated Nunc (ThermoScientific, UK) plate at a density of 12000cells/cm^2 and incubated for 2 days. After reaching around 80% confluence, RMM+ medium was removed, cells were washed twice with pre-warmed RMM- to remove any residual growth factors, and replaced with RMM- containing 10µM DAPT and 100nM ATRA. Cells were incubated for 48h. This stage is referred to as DAPT/ATRA pre-differentiation stage. After 48h, the medium was changed to RMM- only, and

cells were observed daily, with medium change every 2 days till the end of the experiment. Cells were differentiated for 7 and 14 days, which refers to 7 days or 14 days after the pre-differentiation stage (after DAPT/ATRA were removed and replaced with RMM- medium).

3.3.2.3. Lentiviral infection of SPC-04 cells

One day before transduction, SPC-04 cells were seeded in 6-well laminin-coated Nunc plates at a density of 20000 cells/cm^2. The following day, growth media was removed and replaced with RMM- containing 10µM DAPT and 100nM ATRA and a concentration of $3.5x10^5$ TU/ml of lentivirus per well. The pre-differentiation stage was chosen for lentiviral infection based on previous observations showing higher cell survival compared to viral transduction at undifferentiated stage. Cells containing lentiviral particles were incubated for 48h at 37°C, 5% CO_2. Transduction efficiency was assessed using a fluorescent microsope (Leica DMIL) equipped by a Leica camera DFC420C (x10 objective). The cells were lysed for RNA extractions using the QIAgen RNeasy Mini Kit (QIAGEN, UK), followed by cDNA synthesis (as described in 3.3.1.12 and 3.3.1.13 respectively).

3.3.2.4. Quantification of MYT1L levels in SPC-04 cells

qPCR reaction was performed as described in section 3.3.1.14. *RPL18* (60S Ribosomal Protein L18) was used as a housekeeping gene for normalisation of SPC-04 cDNA and ΔCt values ($\Delta Ct = Ct_{TARGET} - Ct_{RPL18}$) were produced for analysis. *RPL18* expression was found to be relatively stable across different stages of neural stem cell differentiation. Sequences of the primers used can be found in Table 3.5.

Table 3.5. Primer sequences used for qPCR amplification of *MYT1L* and the control gene *RPL18* in SPC-04 cells.

Primer	Sequence
hRPL18 Forward	5' – GAGAGGTGTACCGGCATTTC – 3'
hRPL18 Reverse	5' – CTCTGGCACGCTCGAACT – 3'
hMYT1L Forward	5' – TGGAGAGCAACCTGAAGACC – 3'
hMYT1L Reverse	5' – ATTCCTCTCACAGCCTGCTT – 3'

ΔCt values for each were calculated using Ct values from one independent experiment. ΔCt value for each condition was subtracted from the ΔCt of the control condition (i.e. cells transduced with non-silencing shRNA) producing ΔΔCt. The changes in mRNA expression relative to control were then presented as a relative fold expression, using the formula $2^{-\Delta\Delta Ct}$.

All qPCR data were analysed by a one-way ANOVA with *MYT1L* dosage and differentiation stage as factors. All statistical analyses were performed using IBM SPSS Statistics 20 software (IBM Corp., USA). A p-value < 0.05 was considered statistically significant. The data was expressed as mean $\pm$ Standard Error of the Mean (S.E.M.).

3.3.3. RNA labelling and microarray hybridisation

The RNA labelling, hybridisation, washing and scanning steps were outsourced and carried out by High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, Oxford. RNA integrity was analysed using Agilent RNA 6000 Nano Kit (Agilent Technologies, Inc; Germany) on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc). Total RNA was amplified and labeled (biotinylation) using the TargetAmp™-Nano Labeling Kit for Illumina Expression BeadChip (Cambio, UK) according to the manufacturer's instruction. A whole genome gene expression profiling was run on Illumina's HumanHT-12 v4.0 Expression BeadChip, which contains over 47000 probes. Samples were randomly allocated across the chip. The hybridised and washed chips were scanned using an Illumina iScan Scanner.

3.3.4. Statistical and bioinformatic analyses of microarray data

Experiments were performed in triplicates for cells differentiated for 14 days, and quadruplets for cells differentiated for 7 days. RNA from each condition was included in the microarray. However, post microarray analyses revealed, that one sample from 14days of differentiation infected with shRNA *MYT1L*, did not meet the required criteria due to low RNA quality, and was excluded from further analysis. Thus, results of differential gene expression at 14 days of differentiation could not be analysed separately (experimental group consisted only of 2 independent samples) and they were combined with results from day 7 of differentiation. Raw data was extracted using GenomeStudio Data Analysis Software, and was further processed by Dr. Sylvane Desrivières in R statistical environment (http://www.r-project.org) using a Lumi package. The variance stabilising transformation method was used, followed by a quantile

normalisation. Probes with the unreliable expression measurements were flagged and removed from the analysis. Next, based on the coefficient of variation, genes whose expression was constant across all the experiments (i.e., was invariant) were removed from the dataset. Of the 47,230 probes that feature on the Illumina platform, data from 9,673 probes were retained for the further analysis. Using the Benjamini and Hochberg FDR method, the p-values were corrected for multiple testing, by generating q-values (Benjamini & Yekutieli, 2001). All differentially expressed transcripts between *Myt1l* knock-down and non-silencing controls with a false discovery rate (FDR) q-values <0.05 were selected for bioinformatic analyses. Transcripts from a total of 170 distinct genes were selected based on the above criteria.

The functional annotation clustering tool, part of the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (http://david.abcc.ncifcrf.gov/), was used to identify enriched biological themes and functional-related gene groups in gene lists generated from microarray data. DAVID systematically maps a large number of genes in a given list to the associated biological annotation terms (i.e. GO terms). Probe IDs from the Illumina HumanHT-12 v4.0 Expression BeadChip array corresponding to each differentially expressed gene were submitted to DAVID, whilst all known transcripts probed by the closest array to this one (HumanHT-12_V3_0_R2_11283641_A) were used as background for analyses. The DAVID defined default annotation databases (i.e., GO terms) were interrogated using the default medium stringency setting and the threshold for the significant enrichment was represented by an enrichment score (es) ≥1.3 ($p<0.05$) (D. W. Huang, Sherman, & Lempicki, 2008). Additional analyses were performed using Ingenuity® systems software (http://www.ingenuity.com/) to identify relationships, mechanisms, functions, and pathways significantly enriched ($p<0.05$) in the dataset. To test if our set of differentially expressed (DE) genes was enriched for any known susceptibility genes for autism or schizophrenia, the Simons Foundation Autism Research Initiative (SFARI) (https://sfari.org) and the Schizophrenia Gene Resource (SZ, http://www.szgene.org) gene lists were downloaded and interrogated using the *phyper* function in R. SFARI is a comprehensive, up-to-date database for all known human genes associated with autism, which currently contains 528 genes. The SZgene database contained 287 genes implicated in schizophrenia by association studies. For both enrichment analyses, the number of overlapping genes was calculated between DE genes and each disease related gene list. These values were then put into the *phyper* function in R, along with the number of DE genes, number of disease related genes, and number of probe sets included in the analysis, to calculate whether there was significant enrichment.

3.4. Results

3.4.1. Generation of lentiviruses enabling MYT1L overexpression

Attempts were made to try to identify potential interacting partners for Myt1l by overexpressing this gene, therefore allowing us to look at protein-protein interactions and help to determine Myt1l functionality. The human *MYT1L* ORF was initially cloned into the pLenti7.3/V5-TOPO® vector (Life Technologies, UK) to enable exogenous expression in cultured cell lines. Firstly, the *MYT1L* ORF was successfully amplified by high-fidelity PCR, which revealed a specific band of the expected size (~3.55kb) (Figure 3.5). Reactions were carried out in two independent experiments and results were visualised on the gel. Samples with the clearest band (second and third sample on the gel corresponding to V5-MYT1L and Native MYT1L respectively) were chosen for ligation into the expression vector, as they showed only single and specific band.

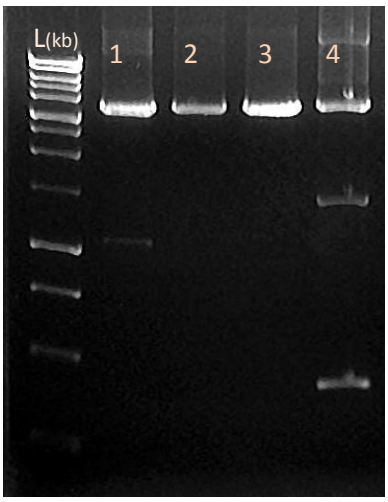


Figure 3.5. Agarose gel electrophoresis to determine amplification of *MYT1L* ORF containing both native and V5-tagged primers. PCR reaction was carried out in duplicates. The first lane, labelled L is the 1kb DNA ladder used as a marker, samples 1 and 2 belong to V5-MYT1L and sample 3 and 4 corresponds to Native-MYT1L. Samples 2 and 3 showed no unspecific binding or primer dimers, and were chosen for ligation into expression vector.

The amplified *MYT1L* insert was successfully ligated into the expression vector in the correct orientation, which was confirmed by HindIII enzyme digest (Figure 3.6). As mentioned in section 3.3.1.9, for sense insertion we expected the following bands sizes: 312bp, 481 bp, 556 bp, 584bp, 2353bp, 2658bp and 3747bp. The bands for V5 (lanes 6 nad 7) and Native (lanes 4 and 5) MYT1L amplicons are of correct sizes for the sense insertion, therefore confirming that the insert was successfully ligated into the expression vector in the correct orientation.

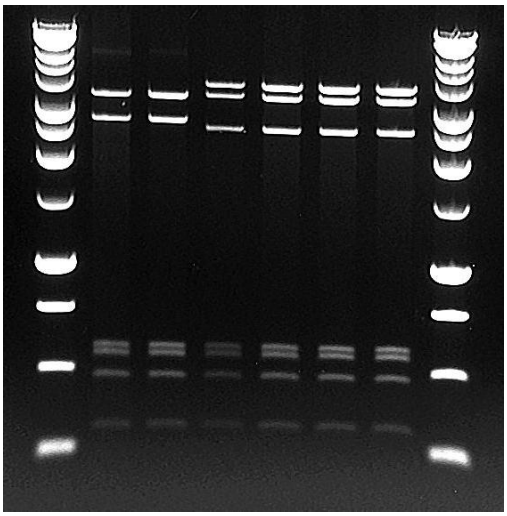


Figure 3.6. Agarose gel electrophoresis of pLenti vector ligated with MYT1L. Lane 2 and 3 represent empty vector, lanes 4 and 5 correspond to pLenti ligated with Native MYT1L, lanes 6 and 7 exemplify vector ligated with V5-MYT1L and lanes 1 and 7 are the 1kb DNA ladder used as a marker.The four smallest bands were common for all samples and pertained to sizes: 312bp, 481bp, 556bp and 584bp. The top bands were specific for either empty vector (2658 bp and 3344bp) or vector with MYT1L insert (2353bp, 2658bp and 2747bp).

The *MYT1L* insert contained within the pLenti vector was then sequenced (data not shown); thereby confirming the absence of any *de novo* mutations that might disrupt MYT1L protein sequence; and verifying that both the V5-tagged and the *MYT1L* native inserts were in the correct frame.

The next step involved production of viruses containing *MYT1L* in either native or V5-tagged form or empty pLenti vector to enable subsequent confirmation of exogenous MYT1L protein expression. For this purpose 293FT cells (supplied with pLenti7.3/V5-TOPO® kit) were transfected with either V5-*MYT1L* or Native-*MYT1L* containing pLenti or empty pLenti vector. The transfection efficiencies of these cell lines were checked indirectly using a GFP tag embedded in the vector, which after numerous experiments indicated an estimated efficiency of ~50 - 60% (Figure 3.7).

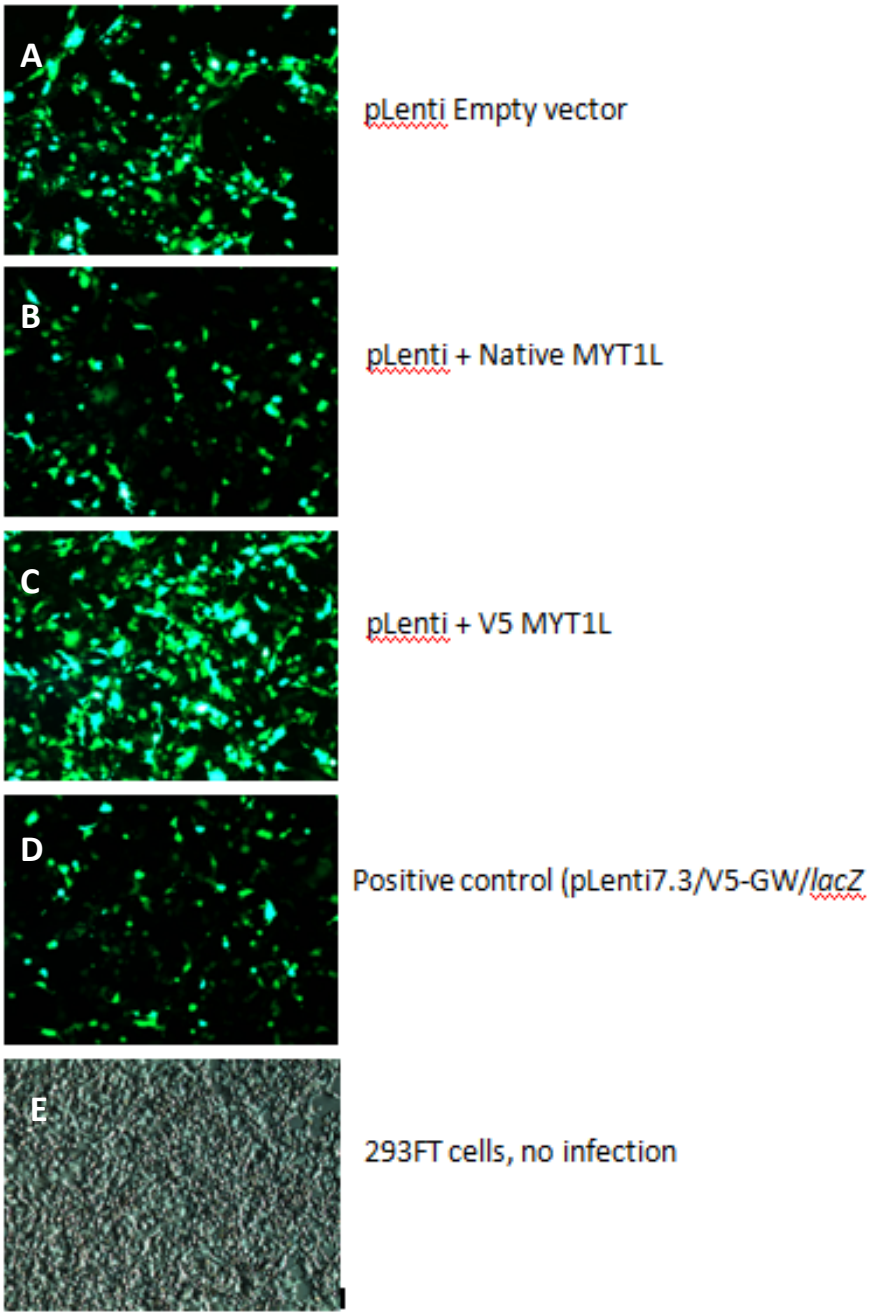


Figure 3.7. The fluorescent microscope images illustrating the GFP expression in 293FT cells obtained after transfection with pLenti empty vector (A), pLenti + Native-MYT1L (B), pLenti plus V5-MYT1L (C), positive control (d) and non infected cells. All photographs were taken using objective X10.

We have transfected a vector supplied with the kit (pLenti7.3/V5-GW/*lacZ*) expressing β-galactosidase to use as a positive control. The positive control yielded a very similar amount of green fluorescent cells, as all the other samples, suggesting there were no problems with the cell transfection. Next, 293T cells were transduced with the virus. However, we only observed

single green cells. Having optimised the protocol several times, by increasing viral particles, longer incubation times with particles, there was no improvement in infection efficiency. As we did not measure viral titers for those cells, one possibility is that the titer was inefficient. Other possible explanations will be mentioned in the discussion.

Since, there was a possibility that virus was successfully transduced but GFP-tag was not visualised, cells were lysed and total RNA extracted and used for qPCR to check mRNA levels of *MYT1L*. Firstly, we checked for genomic plasmid contaminations of our samples, which could mean that any apparent induction of *MYT1L* mRNA expression observed may have been caused by amplification of plasmid DNA contaminants in total RNA samples, as opposed to the corresponding cDNA. To achieve this, a reverse transcription reaction was set-up, where the reverse transcriptase enzyme (Super Script III,Invotrogen, UK) was not included, preventing cDNA synthesis. Thus, any signal obtained from the amplification reaction would reflect only plasmid DNA contamination. Initial data obtained from reaction with reverse transcriptase indicate induction of *MYT1L* (Table 3.6). Considering that lower ΔCt values indicate higher levels of expression, these data indicated a 317-fold change for Native and 595-fold induction for V5 *MYT1L*. However, when control reaction (no reverse transcriptase) was performed we also observed low ΔCt values, indicating that some contaminating plasmid was present in the cDNA samples. When we accounted for the plasmid contamination, we still observed an increase in MYT1L expression (last column of Table 3.6). However, the induction observed was very modest, with only 10-fold change for Native MYT1L and 2.5-fold change for V5. Also, the changes could be at least partially attributed to the endogenous expression of *MYT1L*.

Table 3.6. Δ Ct values from reactions with and without reverse transcriptase to determine amplification of plasmid DNA contaminants in total RNA samples. By excluding reverse transcriptase (SuperScript III) cDNA synthesis is prevented, and all amplification would be due to plasmid contamination with genomic DNA. Data generated this way shows very similar Ct values to ones obtained from cDNA synthesis**.**

Treatment	ΔCt with Superscript III	ΔCt no Superscript III	True ΔCt changes
Empty vector pLenti	7.12	-4.19	11.31
V5	-1.19	-11.20	10.01
Native	-2.10	-10.06	7.96

Although, 293T cells are widely used for transduction experiments, we wanted to verify their efficiency at infecting. Another lab member used these cells to test for overexpression with a

different construct and obtained high levels of exogenous expression, therefore eliminating the possibility of a 'faulty' cell line.

This again suggests that only endogenous *MYT1L* was present in all conditions and we failed to overexpress this gene. Despite several protocol optimisation techniques, such as changing plasmid DNA:Lipofectamine 2000 ratios or increasing the total amount of the DNA:Lipofectamine 2000 complex or leaving viral supernatants with the cells for an additional 24h, the above results persisted, indicating the lack of lentiviral stock containing *MYT1L*-ORF. Considering the lack of success with pLenti and the ViraPower Expression System, we decided to use another commercially available method to assess if the construct itself was problematic. We cloned the MYT1L ORF into the pFLAG-CMV™-4 vector (Sigma, UK) expressing a FLAG tag, which acted in a similar way to the V5-tag, by enabling immunoprecipitation of MYT1L protein using an anti-FLAG antibody, however, similarly to pLenti vector, we were not successful at producing transduced cells, i.e. we observed no green cells present under microscope, and qPCR failed to reveal any changes on mRNA level. As a last resort, we tried In-Fusion® HD Cloning Plus Kit (Clonetech, UK), which allows cloning of any PCR fragment into any linearised vector. We designed primers as per the manufacurer's manual and the expression vector of choice was Lego- iG2 (Lentiviral Gene Ontology (LeGO) vector). This vector was previously used and tested, producing satisfactory results by another laboratory member. After numerous experiments, MYT1L ORF, containing primers for V5 or native form and appropriate restriction site (BamHI in our case), was successfully ligated with Lego-iG2 using the In-Fusion Kit. Unfortunately, subsequent bacterial transformations did not produce any colonies, thus prohibiting cell transfection. The above experiments indicate that although the transfection of *MYT1L* ORF-containing pLenti vectors into 293FT cells was successful, we were unable to detect GFP-positive cells after transduction and mRNA analysis showed plasmid contamination and very modest induction of *MYT1L*.

3.4.2. Use of SPC-04 cells to validate *MYT1L* knock-down

The virus containing shRNA against MYT1L was kindly provided by Lourdes Martinez-Medina (L.M.M). A 3.50 x10^5TU/ml concentrated stock of lentiviruses containing *MYT1L* or non-silencing shRNA was used by L.M.M, which resulted in a 60% reduction of *MYT1L* in SPC-04 cells. Our experiments were designed to test possible gene expression changes in differentiating stem cells upon *MYT1L* silencing.

The human neural stem cell line SPC-04 was used as a neural development *in vitro* model. The cells were transduced with viruses at the pre-differentiation stage (viruses added together with ATRA and DAPT). This allowed us to examine any potential effects of *MYT1L* knock-down on stem cells differentiation patterns. We infected SPC-04 cells with virus expressing either non-silencing shRNA or shRNA targeting *MYT1L.* Firstly, we observed non-infected SPC-04 cell morphology to see if visible phenotypic changes are present upon introduction of *MYT1L* silencing virus. We know that SPC-04 cells can readily differentiate into neurons (Cocks et al., 2013), so any changes in the morphology would have been attributed to the effects of silencing *MYT1L*. Cells showed signs of axonal growth 7 days after differentiation, with more elongated and pronounced morphology being observed 14 days after differentiation. As a control, cells infected with non-silencing shRNA were used, as well as non-infected cells. Cells were collected at 7 days and 14 days of differentiation (after removal of the virus and DAPT/ATRA from medium). Prior to lysate collection for RNA extraction, pictures were taken to assess morphological features of cells infected with the lentiviruses (Figure 3.8). The results show that the morphology of the cells in all experimental conditions changed with differentiation, with cell bodies presenting more elongated forms at day seven (Figure 3.8B) and axonal growth is even more prolonged at day 14 (Figure 3.8E), however there was no visible effects of lentiviral infection on the cells (Figure 3.8C-D and Figure 3.8F-G) confirming previous results from the lab, that *MYT1L* has no evident effect on phenotypic changes of differentiating stem cells. The efficiency of *MYT1L* knock-down was assessed by qPCR analysis.

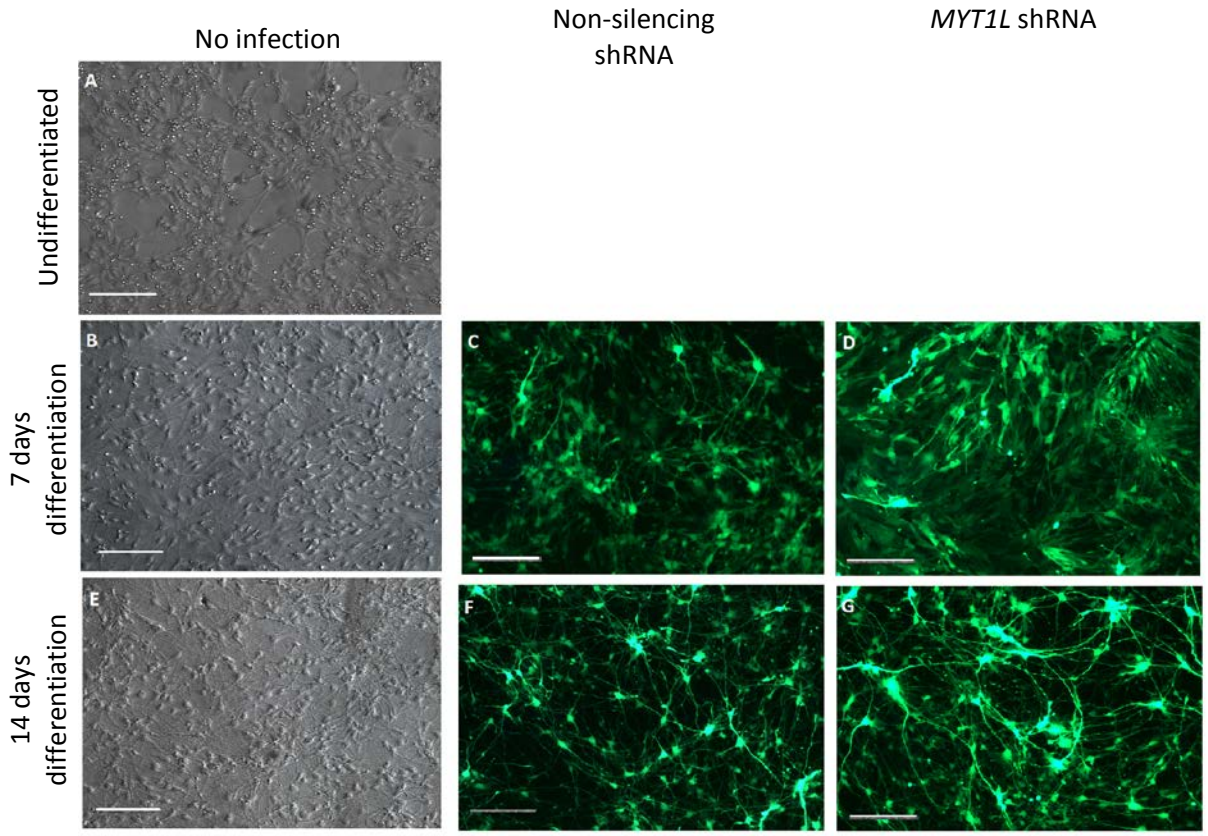


Figure 3.8. Figure representing the human neural stem/progenitor cell line SPC-04 morphology and viral infection efficiency during differentiation and transduction with lentiviral pGIPZ shRNA *MYT1L* and non-silencing. The figures represent cells without the infection (Figure 3.8 A, B and E), cells transduced with non-silencing shRNA (Figure 3.4 C and F) and infection with *MYT1L* shRNA (Figure 3.8 D and G) at two different time points (7 and 14 days. First column shows phase contrast pictures of uninfected cells, second column represents fluorescent images of GFP-expressing cells treated with non-silencing shRNA and last column represents fluorescent images of GFP-expressing cells with shRNA targeting *MYT1L.* Bar represents 200 μm*.*

One-way ANOVA indicated that MYT1L was significantly reduced by 90% at day 7 of differentiation ($F_{1,6}$=420.32, p=8.76E-07) and by 75% in SPC-04 differentiated for 14 days ($F_{1,4}$= 55.69, p=0.002) (Figure 3.9). On the other hand, we observed no significant difference in MYT1L expression between day 7 and day 14 of differentiation ($F_{1,8}$ = 0.018, p = 0.895). Inspection of the dissociation curves for *MYT1L* expression and *RPL18* as a housekeeping gene in qPCR experiments revealed a single amplicon for each measurement, confirming primers specificity (Figure 3.10).

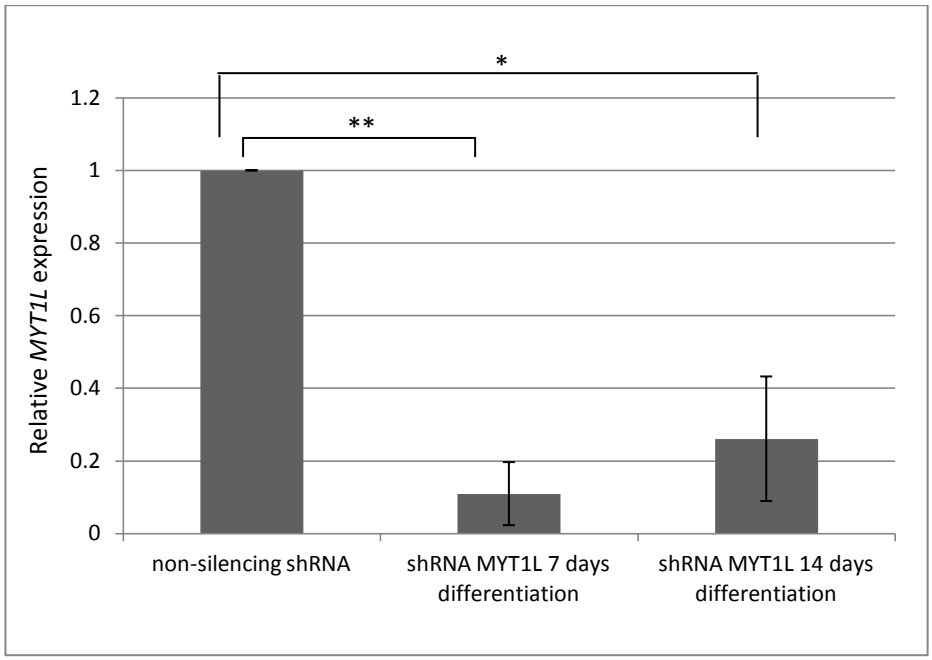


Figure 3.9. Figure showing reduced *MYT1L* mRNA levels in SPC-04 cells infected with *MYT1L* silencing shRNA compared to cells infected with non-silencing shRNA. Quantification of mRNA levels was performed by real-time qPCR, of at least 3 independent experiments, error bars represent +/- 1SD from the mean. *MYT1L* was significantly knocked down in cells at either 7 days or 14 days post differentiation, * - p=0.002, ** - p=8.76E-07

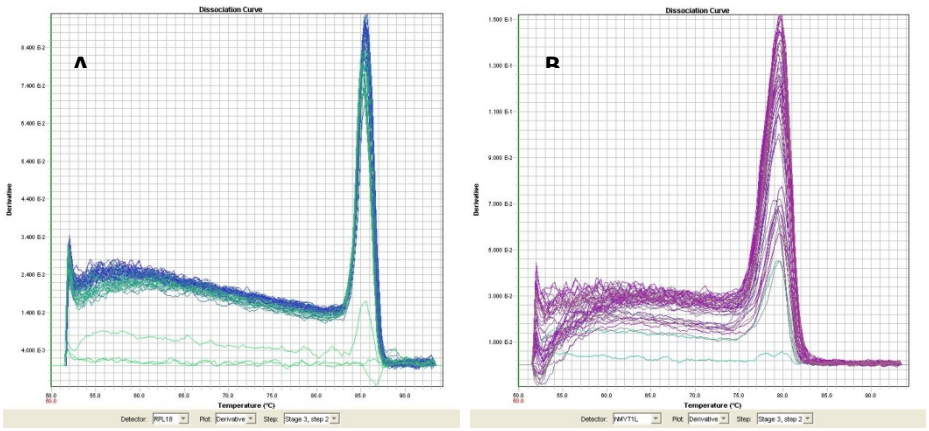


Figure 3.10. Confirmation of *RPL18* (A) and *MYT1L* (B) qPCR primer specificity. Dissociation curve from qPCR analysis of *RPL18* and *MYT1L* indicated a single, distinctive peak (i.e. one PCR product), indicating specificity of the primers used.

3.4.3. MYT1L knock-down alters expression of other genes during SPC-04 differentiation.

After verifying *MYT1L* knock-down efficiency during stem cells differentiation, microarray analyses were performed in order to identify potential downstream transcriptional targets of *MYT1L.* This allowed us to compare gene expression profiles obtained from SPC-04 cells infected with *MYT1L* shRNA differentiated for 7 and 14 days to those of non-silencing shRNA, which acted as controls.

A total of 170 illumina probes were differentially expressed (FDR < 0.05) between *MYT1L* silencing and non-silencing cells, which corresponded to transcripts from 141 distinct genes (105 downregulated in *MYT1L* knock-down cells, 36 up-regulated). These finding can suggest that MYT1L can regulate the expression of those genes. The list of the top 30 differentially expressed (DE) genes is presented below (Table 3.7), and a full list of DE genes is attached as an appendix (Table A1). Unfortunately, the expression of *MYT1L* did not reach the detection threshold at any differentiation time point or condition; hence its downregulation could not be corroborated by microarray. This could be due to the fact that *MYT1L* expression, although detectable by qPCR, in differentiating stem cells, is generally low, and therefore could have been below detection levels set by microarray experiments. This will be discussed in more details later.

Table 3.7. List of top 30 differentially expressed (FDR < 0.05) genes between *MYT1L* knock-down and non-silencing shRNA controls in SPC-04 cells. Relevant illumina probe, gene symbol and a name, log fold change (FC) for cells differentiated for 7 and 14 days, as well as uncorrected and FDR-corrected *p*-values are listed for each gene. Genes are sorted according to *p*- value; lowest to highest.

Probe ID	Gene Symbol	Definition	MYT1L day7 LogFC	MYT1L day14 Log FC	P-Value	FDR
520474	ODZ4	odz, odd Oz/ten-m homolog 4 (Drosophila) (ODZ4), mRNA.	-1.65	-1.97	8.27E-07	0.003904
3450138	CTSC	cathepsin C (CTSC), transcript variant 1, mRNA.	1.66	0.80	8.90E-07	0.003904
5490019	GPX3	glutathione peroxidase 3 (plasma) (GPX3), mRNA.	-0.75	-1.07	1.38E-06	0.003904
5270367	CTSC	cathepsin C (CTSC), transcript variant 1, mRNA.	1.48	0.49	2.01E-06	0.003904
4220068	CD70	CD70 molecule (CD70), mRNA.	1.36	1.11	2.33E-06	0.003904
4890181	RAP1GAP	RAP1 GTPase activating protein (RAP1GAP), mRNA.	-1.40	-2.24	2.42E-06	0.003904
4290403	CMTM7	CKLF-like MARVEL transmembrane domain containing 7 (CMTM7), transcript variant 2, mRNA.	0.86	0.80	4.63E-06	0.006394
4050025	0	mRNA; cDNA DKFZp686J23256 (from clone DKFZp686J23256)	-2.03	-1.67	7.58E-06	0.008798
6620008	KAL1	Kallmann syndrome 1 sequence (KAL1), mRNA.	-2.17	-1.94	8.19E-06	0.008798
3930026	RASL10A	RAS-like, family 10, member A (RASL10A), transcript variant 2, mRNA.	-1.02	-2.04	9.52E-06	0.00921
6250019	SHC1	SHC (Src homology 2 domain containing) transforming protein 1 (SHC1), transcript variant 2, mRNA.	1.11	1.25	1.35E-05	0.01186
670386	ID1	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein (ID1), transcript variant 2, mRNA.	1.27	1.67	1.82E-05	0.012904
2100446	BCAN	brevican (BCAN), transcript variant 1, mRNA.	-3.19	-2.53	1.83E-05	0.012904
3370162	SPON1	spondin 1, extracellular matrix protein (SPON1), mRNA.	-0.39	-1.43	1.96E-05	0.012904
7050575	PPAP2B	phosphatidic acid phosphatase type 2B (PPAP2B), transcript variant 1, mRNA.	-1.52	-1.07	2.00E-05	0.012904
2570564	HLA-DRA	major histocompatibility complex, class II, DR alpha (HLA-DRA), mRNA.	-0.58	-1.82	2.37E-05	0.014347
6110152	C21orf62	chromosome 21 open reading frame 62 (C21orf62), mRNA.	-1.07	-1.51	2.52E-05	0.014357
110333	LOC440585	PREDICTED: hypothetical LOC440585, transcript variant 3 (LOC440585), mRNA.	-1.76	-1.67	2.93E-05	0.015223
4060433	MAOB	monoamine oxidase B (MAOB), nuclear gene encoding mitochondrial protein, mRNA.	-1.74	-2.05	3.17E-05	0.015223
3310307	C21orf63	chromosome 21 open reading frame 63 (C21orf63), mRNA.	-1.45	-1.71	3.27E-05	0.015223
3780092	TNFRSF21	tumor necrosis factor receptor superfamily, member 21 (TNFRSF21), mRNA.	-1.33	-0.70	3.30E-05	0.015223
1410403	LRRN2	leucine rich repeat neuronal 2 (LRRN2), transcript variant 2, mRNA.	-0.41	-0.92	4.35E-05	0.019105
4890021	36403	septin 3 (SEPT3), transcript variant B, mRNA.	-1.72	-1.51	5.23E-05	0.019215
7050333	HEPACAM	hepatocyte cell adhesion molecule (HEPACAM), mRNA.	-0.68	-1.56	5.28E-05	0.019215
2600465	SH3GL2	SH3-domain GRB2-like 2 (SH3GL2), mRNA.	-0.97	-1.08	5.38E-05	0.019215
7000577	GYPC	glycophorin C (Gerbich blood group) (GYPC), transcript variant 2, mRNA.	0.78	1.39	5.60E-05	0.019215
1690403	PMP2	peripheral myelin protein 2 (PMP2), mRNA.	-0.36	-1.40	5.93E-05	0.019215
6620538	UBL3	ubiquitin-like 3 (UBL3), mRNA.	-0.96	-1.07	6.04E-05	0.019215
2760239	RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated) (RASGRP1), mRNA.	-0.95	-1.10	6.05E-05	0.019215
6040451	SLC47A2	solute carrier family 47, member 2 (SLC47A2), transcript variant 1, mRNA.	-0.15	-1.15	6.29E-05	0.019215

3.4.3.1. Functional annotation clustering DAVID of differentially expressed genes

To determine whether this set of differentially expressed genes was enriched for any functional terms, the functional annotation-clustering tool, part of the DAVID (v6.7) bioinformatic resource (D. W. Huang et al., 2008) was employed. DAVID is a web-based bioinformatics application that identifies enriched biology associated with large gene lists derived from high-throughput genomic experiments, such as microarray. Upregulated and down-regulated genes were analysed separately, as this provides an insight into a more specific functional clustering based on the similarity of gene expression pattern. Of the 170 Illumina probes submitted to DAVID, 142 (36 up-regulated and 105 down-regulated) were recognised as probing distinct genes in the human genome. A total of 42 functional gene clusters (16 for up-regulated genes, and 26 for down-regulated genes) were generated from these sets of DE genes. One cluster from up-regulated gene list achieved an enrichment score (es) ≥ 1.3 ($p < 0.05$). The enrichment score provides an indication of the biological significance of the gene groups being analysed. Briefly, the top scoring cluster (ES= 2.04) consisted of genes enriched for vesicle-mediated transport, membrane organisation and homeostasis (Figure 3.11).

DAVID functional-annotation clustering analysis of 105 down-regulated genes produced 2 significantly enriched (es > 1.3) functional clusters. To better understand the details of enriched annotation terms associated with this gene list, the functional annotation chart option was invoked. The enriched term and its associated statistical values is listed below (Table 3.8). Briefly, the only significant GO-term (GO:0007155; Bonferroni corrected p= 0.0453) was cell adhesion, and it consisted of 14 genes.

Figure 3.11. 2-D view modules from DAVID illustrating functional gene cluster significantly enriched (es=1.97) within the list of up-regulated genes differentially expressed between SPC04 cells infected with silencing *MYT1L* and non-silencing controls. The cluster pertain to transport, membrane organisation, endocytosis homeostasis.

Table 3.8. The most enriched annotation term from DAVID chart performed on genes that were downregulated in *MYT1L* knock-down SPC-04 cells. The enrichment p-value and gene count and percentages of total number of genes submitted for analysis are listed along the gene names.

GO-Term	Name	Count	% of total no of genes	p-value	Genes	Bonferroni
GO:0007155	cell adhesion	14	13.33	4.9E-05	LRRN2, NELL2, ASTN1, CD99, BCAN, NLGN3, MEGF10, VCAM1, HEPACAM, HEPN1, KAL1, AGT, GPR56, NCAN, SPON1	0.045

3.4.4. Ingenuity pathway analysis (IPA).

Further analyses were then performed to determine whether the set of DE expressed genes obtained from microarray data was enriched in any biological pathways or diseases as defined by IPA (Ingenuity System Inc, USA). The 'Core Analysis' function included in IPA was used to interpret the data in the context of biological processes, pathways and diseases. The genes that showed significant DE at FDR corrected p-value <0.05 were used for this analysis. IPA mapped these genes to annotated loci within the Ingenuity Knowledge database, where >95%

loci were mapped. Here, for the downregulated set of DE genes, the top associated biological function highlighted by IPA was nervous system development and function, as 9 molecules in this network had an altered expression in *MYT1L* knock-down stem cells (Table 3.9). The top associated diseases and disorders were hereditary disorders and neurological diseases (Rett syndrome, p=5.36E-06). The most common network for downregulated genes was cellular movement, cardiovascular system development and function, cell to cell signalling and interaction (Figure 3.12). These results are in line with what we found using DAVID, and they provided additional information about specific diseases and pathways that were associated with DE expressed genes from *MYT1L* knock-down dataset. For the upregulated set of genes, the most common network highlighted by IPA was cell signalling, small molecule biochemistry and cellular movement (Figure 3.13).

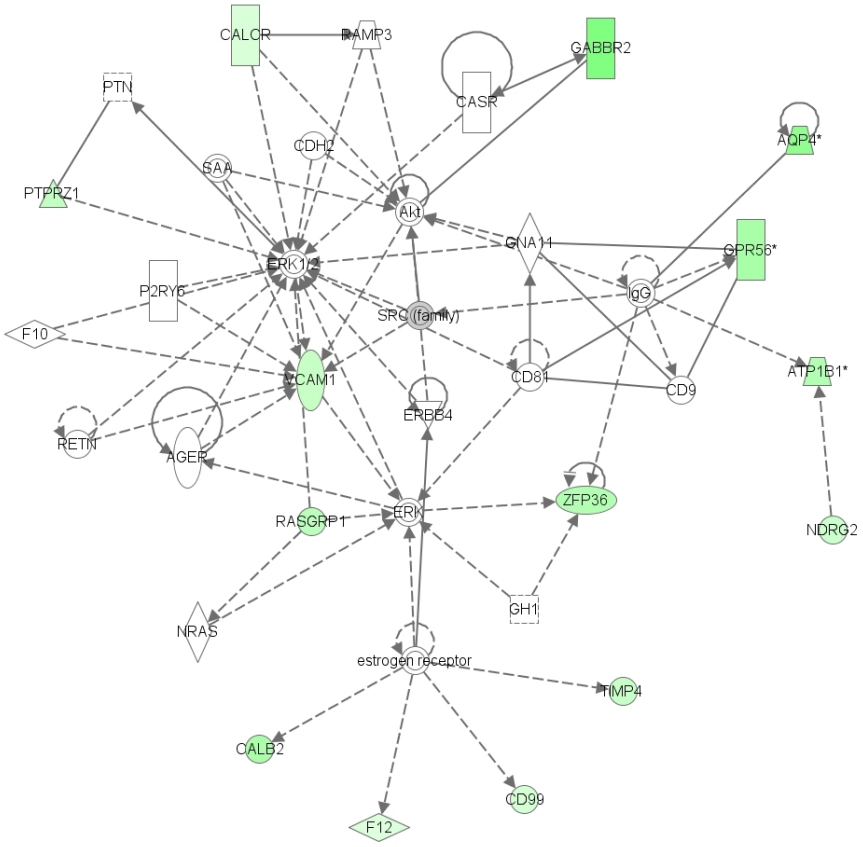


Figure 3.12. Top network in the downregualted set of DE genes in MYT1L knock-down stem cells. Cellular movement, cardiovascular system development and function, cell to cell signalling and interaction. Green indicated genes which were significantly down-regulated in the MYT1L knock-down cells, compared to controls.

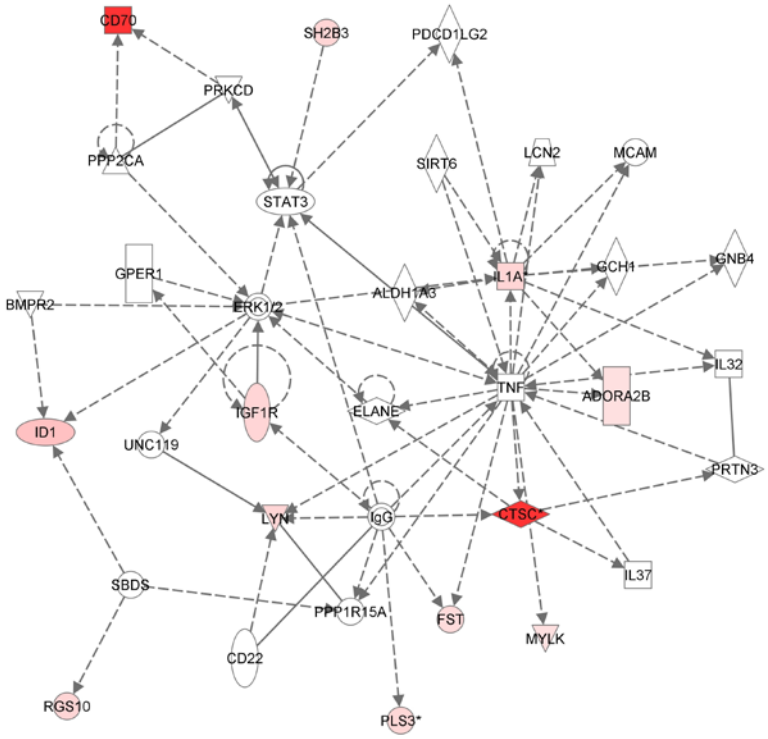


Figure 3.13. Top network in the upregulated set of DE genes in MYT1L knock-down stem cells. Cellular signalling, small biochemistry and cellular movement. Red indicated genes which were significantly up-regulated in the MYT1L knock-down cells, compared to controls

Table 3.9. Ingenuity analyses showing top 20 diseases, disorders, and molecular and cellular functions with significant ($p < 0.05$) enrichment of genes downgulated in *MYT1L* knock-down SPC-04 cells. The corresponding *p*-value and relevant gene symbols are also included.

Category	Function	Function Annotation	p-value	Molecules
Nervous System Development and Function	development	development of central nervous system	9.50E-04	GPR56, PHGDH, PROX1, PTPRZ1, SH3GL2
Hereditary Disorder	Rett syndrome	Rett Syndrome	5.36E-03	ATP1B1, BEX4, GPR56, NELL2
Neurological Disease	Rett syndrome	Rett Syndrome	5.36E-03	ATP1B1, BEX4, GPR56, NELL2
Nervous System Development and Function	development	development of dorsal spinal cord	6.56E-03	PROX1
Nervous System Development and Function	maturation	maturation of neuronal progenitor cells	6.56E-03	WNT7A
Nervous System Development and Function	quantity	quantity of neurosphere cells	6.56E-03	WNT7A
Amino Acid Metabolism	transport	transport of glycine	6.56E-03	SLC6A9
Cardiovascular System Development and Function	chemotaxis	chemotaxis of lymphatic endothelial cells	6.56E-03	PROX1
Cardiovascular System Development and Function	vasodilation	vasodilation of kidney	6.56E-03	AGT
Cell Cycle	cell cycle progression	cell cycle progression of adipoblasts	6.56E-03	AGT
Cell-To-Cell Signaling and Interaction	immune response	immune response of tumor cell lines	6.56E-03	AQP4
Cell-To-Cell Signaling and Interaction	interaction	interaction of dendritic cells	6.56E-03	VCAM1
Cell-To-Cell Signaling and Interaction	response	response of brain cancer cell lines	6.56E-03	AQP4
Cellular Development	maturation	maturation of neuronal progenitor cells	6.56E-03	WNT7A
Cellular Development	quantity	quantity of neurosphere cells	6.56E-03	WNT7A
Cellular Growth and Proliferation	quantity	quantity of neurosphere cells	6.56E-03	WNT7A
Cellular Movement	chemotaxis	chemotaxis of lymphatic endothelial cells	6.56E-03	PROX1
Connective Tissue Development and Function	cell cycle progression	cell cycle progression of adipoblasts	6.56E-03	AGT
Embryonic Development	formation	formation of lens placode	6.56E-03	PROX1
Embryonic Development	quantity	quantity of neurosphere cells	6.56E-03	WNT7A

3.4.5. Enrichment for SFARI and SZ genes

To further elucidate the potential relevance of deregulated genes for neurodevelopmental disorders, DE genes were analysed for enrichment of genes that have previously been associated with ASD or schizophrenia (Table 3.10). We have found a small number of overlapping genes for both disorders; however, we did not see significant enrichment for either set of disease related genes.

Table 3.10. Enrichment of known ASD and SZ genes in the DE data set. The number of overlapping genes between the SFARI or SZ datasets and the DE genes lists, and the p values.

SFARI (no. of overlapping genes)	SFARI p-value	SZ (no. of overlapping genes)	SZ p-value
6 (*ASS1, NFIA, SH3KPB1, NLGN3, PLCD1, HEPACAM)*	0.529	4 *(IMPA2, PTPRZ1, MEGF10, KCNN3)*	0.501

3.5. Discussion, study limitations and future direction

The use of human neural progenitors is an important asset for uncovering the genomic patterns of expression in the normal neuronal differentiation and for the study of neurodevelopmental or neuropsychiatric disorders.

Previous studies suggested that MYT1L plays a key role in neuronal development (Jiang et al., 1996; J. G. Kim et al., 1997; Weiner & Chun, 1997). To explore the molecular effect that MYT1L silencing may have at the level of the transcriptome in differentiating neuronal progenitor cells, a gene expression project was carried out to look at genome-wide expression profiles in the SPC-04 cells infected with shRNA against *MYT1L*.

We identified 170 differentially expressed transcripts, suggesting that signalling via MYT1L protein is involved in the regulation of their transcription rates. The majority of the DE genes were downregulated in the knock-down sample compared to control, suggesting that their expression is regulated by *MYT1L* under vehicle conditions and they can be potential downstream targets of *MYT1L*. The 3 most significant hits were: *ODZ4, RAP1GAP* and *GPX3*. ODZ4 encodes a human homolog of the *Drosophila* pair-rule gene *ten-m* (*odz*) and it is highly expressed in CNS. This gene has been associated with bipolar disorder (Psychiatric, 2011). It has been suggested that it plays a critical role in the myelination of small-diameter axons in the CNS and acts as a regulator of oligodendrocyte differentiation (Suzuki et al., 2012). RAP1GAP is a Ras-like guanine-nucleotide-binding protein (GNBP) that is involved in a variety of signal-transduction processes, it regulates integrin-mediated cell adhesion and might activate extracellular signal-regulated kinase (Daumke, Weyand, Chakrabarti, Vetter, & Wittinghofer, 2004). It has not been associated with neurodevelopment or neurodevelopmental disorders, however its downregulation been associated with human tumours and tumour progression (Tsygankova et al., 2010). Nevertheless, the microarray data should be always interpreted with caution and effects of selected genes from this type of analysis should be further confirmed by additional tests such as qPCR.

Gene set enrichment analyses performed on the DE genes revealed nervous system development as the most common pathway and biological process and the top diseases are neurological or heredity (Rett syndrome). As the MYT1L gene has previously been associated with neurodevelopmental disorders, the results from the enrichment analysis are very interesting. They suggest a gene expression effect that includes genes associated with neurological diseases and nervous system development.. Our findings demonstrate that *MYT1L* is necessary for neurodevelopmental process process, inducing the expression of gene networks involved in neurite outgrowth, axonal development, synaptic transmission and cell adhesion. Since, *MYT1L* has been implicated in neurodevelopmental disease such as schizophrenia and ASD (Y. Lee et al., 2012; K. J. Meyer et al., 2012; S. J. Stevens et al., 2011), we sought to determine whether there may be shared biological processes or pathways between distinct ASD or SZ risk genes in our dataset. We found several overlapping genes for both ASD and SZ, however their correlation was not significant. This could be simply due to the fact that we lacked power to detect significant results in our dataset. The analysis are based on the number of probes submitted for analysis, thus the higher the number of probes the higher the probability of obtaining significant result.

We have also performed microscopic examination of stem cells, which showed no distinctive morphological differences between cells infected with lentiviruses compared to non-infected cells. Interestingly, loss of *Myt1* function (close *Myt1l* paralog) can be, at least partially, compensated by *Myt1l* activity (S. Wang et al., 2007). Therefore, it is possible that consequences of *MYT1L* inactivation in the differentiating neuronal stem cells could be masked by the activation of its paralogs (such as *MYT1* and *MYT3).* It is also plausible to suggest that *MYT1L* might not have a detrimental or very obvious effect on stem cells differentiation, but instead it has a more subtle influence on changes within individual genes or pathways. The presence of genes involved in synaptic transmission and axonogenesis can suggest involvement of MYT1L in those processes. Interestingly, two genes: *WNT7A* and *NLGN3* were repeatedly present with aforementioned GO terms. WNT7A is a member of the Wnt family of signaling molecules which are involved in axon and dendrite development, and synaptogenesis (Salinas & Zou, 2008). Wnt7a signalling specifically regulates the formation and function of excitatory synapses (L. Ciani et al., 2011) and it has been shown to stimulate the proliferation of neural progenitors derived from embryonic mouse brains (Viti, Gulacsi, & Lillien, 2003). It has been suggested that Wnt7a regulates genes involved in neuronal differentiation (Qu et al., 2013). Additionally, gene encoding a member of the neurolignin family of neuronal cell surface proteins – *NLGN3* has also been downregulated upon ablation of MYT1L. A point mutation of

NLGN3 (R451C) was first reported in two affected brothers, one with autism and the other with Asperger syndrome (Jamain et al., 2003). It has been found that NLGN3 can promote the formation of presynaptic elements in cultured hippocampal neurons and that autism-associated mutation lead to a loss of this activity (Chih, Afridi, Clark, & Scheiffele, 2004). It has been shown that NLGN3 mutation does not alter neuron morphology but it impairs synaptic transmission (Földy, Malenka, & Südhof, 2013). Taken together this may suggest that MYT1L contribute to shaping synapses and their properties without altering cell morphology.

Part of this project was devoted to functionally characterise human MYT1L protein, by generating an overexpression model that would have likely proved useful for this work. The aim was to identify other transcription factors or other proteins with which MYT1L interacts, thereby providing clues as to the biological context and function of MYT1L. To achieve this goal, the human *MYT1L* open-reading frame (ORF) was cloned into an expression vector (pLenti7.3/V5-TOPO). Unfortunately, even though cloning was successful we failed to detect high levels of exogenous *MYT1L*. Although GFP-expressing cells were transfected with over expressiong MYT1L, we failed to produce infected cells. One possible explanation for this is the problem with the purchased plasmids. All kits came with pre-mixed packaging system which could be the source of our problem. Since the virus packaging mix is pre-made, we were unable to separate and verify each of the plasmids. The lack of success with all the system is baffling as the sequencing confirmed absence of frame shifts, *de novo* mutations, or any sequence aberrations that would explain lack of progress. Alternatively, some not yet defined factors that are cell specific could have prevented exogenous *MYT1L* expression. Genomic DNA plasmid contamination was detected in our samples, even after DNase I treatment and PCR purification. The possible solution would be to use different a method of RNA extraction (i.e. using Trizol) or to use a different DNase I. Also, it would have been beneficial to perform antibiotic selection (i.e. puromycin selection) on positive clones, to obtain stably transduced cells that were only expressing MYT1L gene and GFP.

Therefore, at this point, it is difficult to draw any conclusions concerning the MYT1L protein and potential targets of *MYT1L* upon its upregulation compared to controls. It would be interesting to see the effects of exonegously expressed *MYT1L* on undifferentiated stem cells, in which this gene is present at very low levels. Further research should concentrate on trying to characterise MYT1L protein function, beginning with additional attempts to try to identify

potential interacting partners, which in turn can lead to unraveling the specific molecular mechanisms driving the apparent association between *MYT1L* and neural development.

There are a number of limitations that are associated with this study. One limitation comes from microarray analysis, in which we observed no significant evidence (FDR < 0.05) of reduced *MYT1L* mRNA expression in *MYT1L* silenced SPC-04 cells compared to cells infected with non-silencing shRNA, although qPCR analyses demonstrated a significant reduction of this gene in a silencing samples. As mentioned before, one possible explanation for the discordance of results lay within microarray sensitivity. Microarray is a powerful tool for studying the molecular basis of interactions on a scale that is impossible using conventional analysis, but it not without limitations. The sensitivity threshold of microarray measurements defines the concentration range in which accurate measurements can be made. Although microarray sensitivity is impressive, it might still be insufficient to detect relevant changes in low abundance genes, such as transcription factors (Czechowski, Bari, Stitt, Scheible, & Udvardi, 2004; Draghici, Khatri, Eklund, & Szallasi, 2006). Therefore, the ability to detect changes in the expression of specific individual genes might be affected. Thus, we found that *MYT1L* probes were present on the array but they produced low signal from each condition, suggesting that this gene can be below the microarray sensitivity threshold. Another limitation comes from pooling data from day 7 and day 14 of differentiation. We have analysed the combined dataset which could lead to potential errors. Due to pooling we could have missed genes that were differentially expressed at one time point but not the other compared to controls. We were also unable to distinguish or measure the magnitude of difference of DE genes between both time points. Despite the aforementioned issues, at the time of writing this thesis, this is the first study showing effects of MYT1L knock-down in SPC-04 stem cells and investigating its role and potential putative targets by microarrays.

Additionally, there is an issue of validation of our findings by an independent method. It would have been beneficial to test the top hits from microarray study by qPCR to confirm and validate our findings. However, due to time constraints within this thesis this was not performed.

It is also noteworthy that SPC-04 cells are derived from a 12-week old foetus and it is possible that the genes expressed by this line differ from the primary cultures or cells *in vivo*. Despite that, neural stem cells are very good models to study neurodevelopment, especially due to the

possibility to manipulate gene expression at various stages of differentiation in order to investigate gene function.

In terms of future studies, further research is required to determine the role of *MYT1L* in neurodevelopment. Since we did not observe any obvious phenotypic changes associated with *MYT1L* down-regulation, it would be interesting to characterise neuronal properties of the cells with deregulated *MYT1L,* by for example by staining cells with different cell markers, such as a presynaptic marker (synaptophysin), a postsynaptic marker (PSD-95) and a neurite marker (MAP2). The stained neurons can be analysed by fluorescence microscopy and compared to non-infected cells. It would also be beneficial to perform immunostaining in undifferentiated cells using markers such as Sox1 and Nestin as well as differentiation markers such Tuj1 (neurons) and GFAP (glia) to investigate the differentiation potential in this cell line upon Myt1l knock-down and to determine the time period required for producing mature neurons and/or glia.

Additionally, it would have been advantageous to test *MYT1L* downregulation effects *in vivo*. In fact, behavioural tests on mice injected with *Myt1l* silencing virus were performed by another PhD student in our laboratory. Animal experiments can provide a real life functional characterisation of the effects which *Myt1l* downregulation can have on the whole organ or body.

In conclusion, this project has explored the gene expression alterations in differentiating stem cells upon *MYT1L* reduction. The results from this study have supplemented the knowledge of the possible downstream targets of *MYT1L* during brain development. They have indicated the importance of *MYT1L* for neuronal development, and suggest that this transcription factor is involved in the differentiation of human neuronal stem cells.

4. Investigating the putative relationship between MYT1L and MEF2A.

4.1. Introduction

4.1.1. MEF2 transcription factors

Myocyte enhancer factor 2 (MEF2) is a member of the MCM1-agamous-deficiens-serum response factor (MADS) family of transcription factors, identified initially in muscle cells (Yu et al., 1992). Since then, MEF2s have been identified in cells of the immune system where they mediate cell fate decisions (Youn, Sun, Prywes, & Liu, 1999). Also, they are highly expressed in neurons of the central nervous system where they show distinct patterns of expression in different regions of the brain with highest levels in the cerebellum, cerebral cortex and hippocampus (Heidenreich & Linseman, 2004; Lyons, Micales, Schwarz, Martin, & Olson, 1995; She & Mao, 2011).

Through evolution, *C. elegans* and *D. melanogaster* only contain one isoform of MEF2, whereas vertebrates express four different MEF2 isoforms – MEF2A-D. The N-terminus of MEF2 contains a MADS box, which is conserved between species, and it is located adjacent to the MEF2 domain, which is important for dimerisation, DNA binding and co-factor interactions (McKinsey, Zhang, & Olson, 2002). The C-terminal region of MEF2 is the transcription activation domain that can undergo extensive alternate splicing to create complex pattern of gene transcription (McKinsey et al., 2002). MEF2 family members have distinct but overlapping expression patterns in the brain during embryogenesis (Lyons et al., 1995), following that of neuron maturation, which suggests that onset of MEF2 expression coincides with neuronal differentiation.

Recent studies identified the involvement of MEF2 in a number of neurodevelopmental processes, including synapse formation (S. W. Flavell et al., 2006; Pulipparacharuvil et al., 2008). Furthermore, the alterations of downstream target genes of MEF2 have been implicated in a variety of neurological disorders (Greer et al., 2010; Lanz et al., 2013), suggesting that MEF2 is an important transcription factor involved in neurodevelopment.

The transcriptional activity of MEF2 is activity-regulated, in which membrane depolarisation, achieved in cultured neurons using potassium chloride (KCl) stimulation, increases MEF2

transcriptional activity through calcium influx via L-type voltage-gated calcium channels (Mao, Bonni, Xia, Nadal-Vicens, & Greenberg, 1999) and calcineurin activation (Mao & Wiedmann, 1999). During mammalian development, electrical activity promotes the calcium-dependent survival of neurons that have made appropriate synaptic connections. It has been shown that calcium influx into cerebellar neurons triggers the activation of the MKK6-p38 MAP kinase cascade and that the p38 MAP kinase then phosphorylates and activates MEF2s (Mao et al., 1999). Once activated by this calcium-dependent p38 MAP kinase signalling pathway, MEF2 can regulate the expression of genes that are critical for survival of newly differentiated neurons. These findings demonstrate that MEF2 is a calcium-regulated transcription factor that has a well-characterised, defined function during nervous system development that is distinct from previously identified functions of MEF2 during muscle differentiation. Inhibition of MEF2 function in cortical neurons has been shown to cause apoptotic death, implying that MEF2-dependent regulation of transcription is necessary for survival (Mao et al., 1999).

Understanding how MEF2–cofactor interactions are influenced by diverse signalling pathways, both calcium-dependent and independent, represents a major challenge for the future. The fact that MEF2 acts as a transcriptional switch for cellular processes that are disrupted in numerous diseases also makes it an attractive target for pharmacological and genetic modification. For example, the loss of MEF2D-mediated neuronal survival has been shown to underlie the process of the survival of DA neurons in models of Parkinson's disease (Yin et al., 2012) and rare variants in the MEF2C gene have been associated with autism, mental retardation and developmental delay (Neale et al., 2012; Novara et al., 2010).

MEF2 proteins also regulate dendrite morphogenesis, differentiation of post-synaptic structures and excitatory synapse number (S. W. Flavell et al., 2006; Shalizi et al., 2006). Promotion of the post-synaptic differentiation of granule neurons can be achieved through sumoylation, by repressing the expression of the Nur77 transcription factor - a negative regulator of dendritic differentiation (Shalizi et al., 2006). Sumoylation is a post-translational modification that involves the attachment of one or more SUMO groups to a protein and is catalysed by an enzymatic cascade termed the 'sumoylation pathway' (Scheschonka, Tang, & Betz, 2007). A recent study investigated the role of MEF2A in pre-synaptic development in the mammalian brain. Knockdown of *MEF2A* increased the density of orphan pre-synaptic sites in primary neurons and in the cerebellar cortex of rats *in vivo,* while a sumoylated transcriptional repressor form of MEF2A mediated the suppression of orphan pre-synaptic sites (Yamada et al., 2013). Moreover, dephosphorylation of MEF2 by calcineurin regulates the expression of

activity-regulated cytoskeletal-associated protein (*Arc*) and synaptic RAS GTPase-activating protein (*synGAP* or *Syngap1*). Both, ARC and synGAP play important roles in synaptic disassembly by promoting the internalisation of glutamate receptors and by inhibiting Ras-MAP signalling pathway (S. W. Flavell et al., 2006). MEF2A and MEF2D have been reported to negatively regulate synaptic development based on RNAi experiments *in vitro* (S. W. Flavell et al., 2006). However, a more recent study suggested only a subtle role of MEF2A/D in regulating synaptic function (Akhtar et al., 2012). *Mef2a* and *Mef2d* brain-specific double knock-out mice showed only deficits in motor coordination and short-term synaptic plasticity, but did not exhibit any other behavioural changes and had no impact on learning and memory, long-term potentiation or number of synapses (Akhtar et al., 2012). MEF2A and MEF2D have been shown to control expression of genes associated with autism spectrum disorder (ASD) (*Ube3A, Slc9A6, Pcdh10, C3orf58*) (Steven W. Flavell et al., 2008; Lanz et al., 2013; Eric M. Morrow et al., 2008), which led to the proposition that autistic phenotypes can be generated by abnormal regulation of synaptic development and altered MEF2 signalling as a possible cause (Dietrich, 2013).

An understanding of the ways in which signalling pathways are connected to MEF2 targets in the brain remains an interesting and exciting research possibility. Analysis of alterations of expression of such targets in neurodevelopment and neurological diseases might reveal additional information on causation of many disorders in which MEF2 has been implicated.

4.1.2. MEF2 interaction with MYT1L: a preliminary analysis

MEF2 factors activate transcription via binding to A/T rich DNA consensus sequence CTA(A/T)$_4$TAG/A as homo- or heterodimers (Black & Olson, 1998; Cserjesi & Olson, 1991). The ability of MEF2 to regulate neuronal-specific transcriptional programs may occur through DNA-binding site selection. The expression of MEF2 in neurons shows optimal DNA-binding constraints for specific nucleotide sequences that flank the MEF2 site, and this is not observed with MEF2 factors from other cell types (Andrés, Cervera, & Mahdavi, 1995). The additional, brain-specific sequence (5'-TGTTACT(A/T)(A/T)AAATAGA(A/T)-3') was not observed in the skeletal or cardiac muscle extracts.

The analyses of gene expression patterns during stem cells differentiation, performed in our laboratory, have revealed the aforementioned brain-specific MEF2 binding site as present in most genes co-expressed with MYT1L. Briefly, microarray analysis of SPC-04 cells at different

stages of differentiation was performed in experiments consisting of three independent replicates of undifferentiated cells, pre-differentiated cells and cells that had been differentiated for 7 days (Desrivieres et al., 2014). Data analysis was performed by Dr Sylvane Desrivières and Dr Anbarasu Lourdusamy.

Normalised and FDR-corrected results showed 270 genes that were co-expressed with MYT1L at day 7 of differentiation. Many of those genes have been linked to neurogenesis or differentiation. This set of co-expressed genes was significantly enriched for the presence of MEF2 transcription factor binding site (n = 114 genes, fold enrichment = 1.69; FDR-adjusted p-value = 1.40E-09; see Table 4.1). These results showed us that the MEF2 biding site is present somewhere within each of those genes. Being interested in the potential link between MYT1L and MEF2 and wanting to see if the putative MEF2 binding site is present within the promoter portion of the genes of interest, we examined the data further by interrogating sequences 10kb downstream and 2kb upstream of the transcription start site of each of those genes in search of MEF2-binding site. To perform this task, we have used TFBS Conserved table, part of UCSC genome browser, by coping all of the genomic coordinates and customised it to display only MEF2A sites (restricting the list to display only V$MEF2).This way, we have found that 190 of 270 genes, including MYT1L, contain the brain-specific MEF-2 sequence.

Table 4.1. Top 20 transcripts which correlated with the expression of *MYT1L* during differentiation of SPC-04 cells. Highlighted in yellow are genes that contain MEF2 brain-specific binding sites in the proximity of their TSS.

Symbol	Definition	Correlation co-efficient R	Correlation P-value
MYT1L	Homo sapiens myelin transcription factor 1-like (MYT1L), mRNA.	1	2.68E-07
JPH3	Homo sapiens junctophilin 3 (JPH3), mRNA.	0.9552052	1.32E-06
SEZ6L2	Homo sapiens seizure related 6 homolog (mouse)-like 2 (SEZ6L2), transcript variant 2, mRNA.	0.9525285	1.75E-06
CDK5R1	Homo sapiens cyclin-dependent kinase 5, regulatory subunit 1 (p35) (CDK5R1), mRNA.	0.95077276	2.10E-06
LRRC24	Homo sapiens leucine rich repeat containing 24 (LRRC24), mRNA.	0.94956756	2.36E-06
ARHGEF7	Homo sapiens Rho guanine nucleotide exchange factor (GEF) 7 (ARHGEF7), transcript variant 2, mRNA.	0.9484351	2.63E-06
KCNQ2	Homo sapiens potassium voltage-gated channel, KQT-like subfamily, member 2 (KCNQ2), transcript variant 3, mRNA.	0.9483695	2.65E-06
MAP4	Homo sapiens microtubule-associated protein 4 (MAP4), transcript variant 3, mRNA.	0.94808483	2.72E-06
ZNF423	Homo sapiens zinc finger protein 423 (ZNF423), mRNA.	0.9458477	3.35E-06
RTN1	Homo sapiens reticulon 1 (RTN1), transcript variant 1, mRNA.	0.94325316	4.21E-06
PTPRO	Homo sapiens protein tyrosine phosphatase, receptor type, O (PTPRO), transcript variant 5, mRNA.	0.9416315	4.84E-06
SNAP91	Homo sapiens synaptosomal-associated protein, 91kDa homolog (mouse) (SNAP91), mRNA.	0.940563	5.28E-06
PKIA	Homo sapiens protein kinase (cAMP-dependent, catalytic) inhibitor alpha (PKIA), transcript variant 7, mRNA.	0.93742937	6.80E-06
PPP2R2C	Homo sapiens protein phosphatase 2 (formerly 2A), regulatory subunit B, gamma isoform (PPP2R2C), transcript variant 2, mRNA.	0.9357074	7.76E-06
BCL11B	Homo sapiens B-cell CLL/lymphoma 11B (zinc finger protein) (BCL11B), transcript variant 1, mRNA.	0.93507975	8.14E-06
GRIA2	Homo sapiens glutamate receptor, ionotropic, AMPA 2 (GRIA2), mRNA.	0.9347996	8.31E-06
ELMO1	Homo sapiens engulfment and cell motility 1 (ELMO1), transcript variant 1, mRNA.	0.93372506	9.00E-06
SLC17A6	Homo sapiens solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 6 (SLC17A6), mRNA.	0.93345857	9.18E-06
PPP2R2C	Homo sapiens protein phosphatase 2 (formerly 2A), regulatory subunit B, gamma isoform (PPP2R2C), transcript variant 2, mRNA.	0.9319617	1.02E-05
PKIA	Homo sapiens protein kinase (cAMP-dependent, catalytic) inhibitor alpha (PKIA), transcript variant 6, mRNA.	0.9312388	1.08E-05

These analyses provided an interesting opportunity and suggested a link between MYT1L and MEF2. As the primary function of transcription factors it so either activate or repress the expression of target genes by interacting with sequence specific DNA motifs and since most of MYT1L co-expressed genes in this data set contain MEF2 binding site, we hypothesised that MEF2 is indeed a regulator of MYT1L and MYT1L co-expressed genes.

4.2. Aims

Given the above analyses indicating that MYT1L and MEF2 co-expressed genes contain a MEF2 binding site in their promoter region, we hypothesised that MYT1L is a target gene of MEF2A. The present study aimed at investigating the link between MYT1L and MEF2A. To achieve this, we first tested if Mef2a expression is conserved in vertebrates using zebrafish embryos and mouse brains as a model for brain development. To address the question of whether MEF2A is involved in stem cell differentiation and whether it regulates expression and function of MYT1L in this process, a gene knockdown approach was utilised in neural stem cells during the differentiation, followed by microarray gene expression analysis.

4.3. Materials and methods

4.3.1. RNA extraction and cDNA synthesis of mouse and zebrafish samples.

RNA from zebrafish embryos at stages: 18SS, 24hpf, 48hpf and 72hps were extracted according to the protocol described in chapter 2.3.3. Briefly, whole embryo RNA was extracted from pooled (25 embryos) embryos at each developmental stage. Whole brain mouse RNAs were extracted from pools of 5 (for the E10 stage) and 3 (for the E14 stage) embryos. RNA extracted from the frontal cortex was used from later developmental stages (E18, P1week, P1month, P6months). In this case, triplicates from independent brains were analysed for each stage up to 6 months, for which data was obtained from a single brain. Mouse RNA samples extracted from CD1 mouse brains at embryonic day 10 (E10), E14, E18, and at postnatal (P) stages 1 week, 1 month, or 6 months were purchased from AMS Biotechnology (Abingdon, UK). cDNA synthesis was carried out according to the protocol described in 2.3.6.

4.3.2. qPCR analysis of mouse and zebrafish extracts.

Quantitative RT-PCR (qPCR) was performed as described in section 2.3.7. Briefly, cDNA samples were amplified using ABI Prism 7900HT sequence detection system installed with ABI Prism ® SDS 2.1 software (Applied Biosystems) in a final volume of 20 µl containing 2x power SYBR green master mix (Applied Biosystems), 4µl diluted cDNA and 0.07µM of forward and reverse primer (Table 4.2). The PCR consisted of three technical replicates and was performed under

the following thermal cycler conditions: initial enzyme activation step of 95°C for 15mins followed by 40 cycles of 95ºC for 30secs and 59ºC for 30secs. Fluorogenic data was collected at the 59ºC stage. The qPCR reaction was evaluated by dissociation curve analysis to ensure that the amplicons generated were specific. Zebrafish *β-actin* was used as the housekeeping gene for qPCR analysis of cDNA produced from zebrafish RNA extracts while the control gene used for analysis of mouse cDNA was *Gapdh.* Zebrafish primers were designed to target *mef2aa* gene, as this form has been identified as the protein coding gene and it is an orthologs of human *MEF2A*.

Table 4.2. Table showing the oligonucleotide sequences of Mef2a primers for qPCR analysis of zebrafish and mouse samples.

Primer		Sequence
zf-actin	Forward	5' – GCAGAAGGAGATCACATCCCTGGC – 3'
	Reverse	5' – CATTGCCGTCACCTTCACCGTTC – 3'
zf-mef2a	Forward	5' – GGCTCTCCAGGGCTCTCTAT – 3'
	Reverse	5' – AAACCAGATGGGGTTACACG – 3'
m-Gapdh	Forward	5' – TGTTCCTACCCCCAATGTGT – 3'
	Reverse	5' – CCTGCTTCACCACCTTCTTG – 3'
m-Mef2a	Forward	5' – TTACTTCCCCTGGAATGCTG – 3'
	Reverse	5' – GGAGGTGAAATTGGCTCTGA – 3'

The qPCR data was analysed by taking the mean of the Ct values for each of the 3 technical replicates for the housekeeping gene (*zf-actin* or *m-Gapdh*) and for *mef2aa* and *Mef2a* expression respectively. In order to generate ΔCt values, the mean Ct for *mef2aa/Mef2a* expression was subtracted from the mean Ct for *zf-actin/m-Gapdh* ($\Delta Ct = Ct_{target} - Ct_{housekeeping\ gene}$) at each stage. ΔΔCt values were then obtained in order to calculate relative expression. Changes in expression at each stage were relative to E10 for mouse and 18SS stage for zebrafish and were calculated as ΔΔCt ($\Delta Ct - \Delta Ct_{E10\ or\ 18SS}$). ΔΔCt was then converted to relative fold expression using the formula $2^{-\Delta\Delta Ct}$. Data are presented as group means and error bars show the standard error (+/-1.0 SE) from the means. The expression data was analysed by a one-way analysis of variance (ANOVA) with developmental stage as a factor.

4.4. Construction of shRNA targeting human MEF2A

A *MEF2A* targeting shRNA was cloned into a lentiviral gene ontology (LeGO) vector (K. Weber, Mock, Petrowitz, Bartsch, & Fehse, 2010) to generate a *MEF2A* silencing plasmid for *in vitro* studies as follows:

4.4.1.1. Vector preparation

LeGO vectors are third-generation lentiviral plasmids, meaning they do not encode any viral proteins but contain *cis*-active elements for packaging, reverse transcription, and integration (Kristoffer Weber, Bartsch, Stocking, & Fehse, 2008). These lentiviral expression vectors efficiently transduce slowly dividing cells, including hematopoietic stem-progenitor cells (HSCs), resulting in stable gene transfer and expression. Additionally, LeGO vectors allow incorporating up to 9kb of foreign sequences. LeGO-G expresses enhanced green fluorescent protein (eGFP), as a marker gene to ensure efficient detection (Figure 4.1).

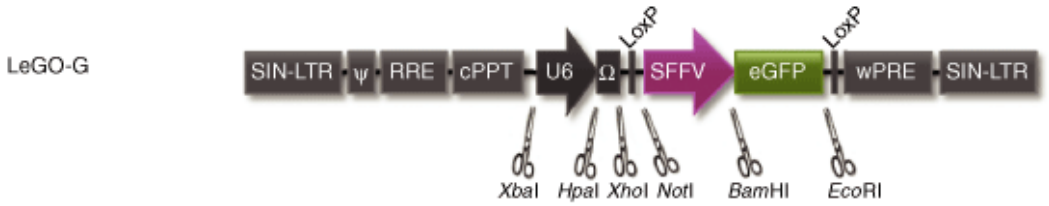


Figure 4.1. The lentiviral gene ontology (LeGO-G) vector principle. LeGO-G vector was used for gene silencing. LeGO-G also contains a retroviral enhancer/promoter of spleen focus-forming virus (SFFV), which allows a broad and high expression pattern for expression of the reporter gene eGFP (enhamnced Green Flouresecent Protein) while the U6 promoter drives the expression of the shRNA cloned into the Multiple cloning site (MCS) using the *HpaI* and *XhoI* restriction sites. Other features include rev-responsive element (RRE); selfinactivating- long-terminal repeat (SIN-LTR); Woodchuck hepatitis virus post-transcriptional regulatory element (wPRE) and central polypurine tract (cPPT). Picture adapted from (Kristoffer Weber et al., 2008).

To linearise the vector, LeGO-G was digested by using the *HpaI* and *XhoI* restriction enzymes. Enzymes and buffers were purchased from NEB, UK. A typical reaction contained 2U of each enzyme, 1µg DNA (vector), 1x Buffer *EcoRI*, 1x BSA made up to a final volume of 20µl with dH_20. The reaction mix was incubated at 37°C for 1 hour to allow complete digestion, followed by 20 minutes at 65°C to inactivate the restriction enzymes. To prevent re-circularisation, the vector was dephosphorylised with 5 units of Antarctic phosphatise, 1µg of digested vector and 1x Antarctic phosphatase buffer and incubated at 37°C for 30 minutes followed by enzyme inactivation for 5 mins at 65°C. Plasmid DNA was purified using the QIAquick PCR purification

kit (QIAGEN, Crawley, UK), according to the manufacturer's instructions and 200ng of the purified, digested and dephosphorylated vector was electrophorised on the gel, to check for complete digestion.

4.4.1.2. MEF2A shRNA oligos

The shRNA against human MEF2A was designed based on previously described shRNA targeting rat Mef2a. We adapted the sequence to target only the human isoform of MEF2A. Target sense sequence is presented in Table 4.3.

To design a hairpin we followed a protocol previously described by Dr Tyler Jacks at MIT (http://www.addgene.org/static/data/94/67/16242780-af64-11e0-90fe-003048dd6500.pdf). Briefly, an *HpaI* site in Lego-G vector leaves a blunt end prior to the –1 position in the promoter. The oligo design incorporated a 5' T in order to reconstitute the –1 nucleotide of the U6 promoter. Therefore the oligo format is as follows:

Sense oligo: 5'T-N19-loop sequence (TTCAAGAGA)-corresponding antisense sequence (91NC)-termination sequence (TTTTTTC)

Anti-sense oligo: complement of sense with additional nucleotides at 5' end to generate *XhoI* Overhang.

To properly assess knockdown, the gene expression level from the scramble control vector was used in comparison with the target-specific shRNA transfected samples. We used another gene from the MEF2 family to produce a non-silencing scrambled shRNA (Table 4.3). The sequence of the MEF2D shRNA was altered in a way that it did not recognise any known human gene, therefore should not lead to any alterations on a molecular or cellular level. We have also used empty, ligated LeGO-G vector as another control.

Table 4.3. Sense sequence of shRNA targetting MEF2A and a scrambled MEF2 as a negative control.

Gene	Sense target sequence
MEF2A	5' – GTTATCTCAGGGTTCCAAT – 3'
MEF2 scrambled	5' – GGTCTCATAGGCTCCGAGAG – 3'

The oligos were ordered from Integrated DNA Technologies (Coralville, IA, United States of America), with 5' phosphates and PAGE purified. 60pmoles of sense and antisense oligos were annealed in a buffer containing 100mM K-acetate, 30mM HEPES-KOH pH 7.4, 2mM Mg-acetate by incubating at 95°C 4min, 70°C for 10 mins, followed by a slow temperature decrease (0.1°C/min) to 4°C.

4.4.1.3. Ligation of annealed shRNA into the LeGO-G plasmid vector

DNA ligation reactions were carried out in the molar ratio 1:1 of annealed oligos and linearised, dephosphorylated LeGO-G vector. 30fmol of each component was mixed with 1x T4 DNA ligase buffer and 1U T4 DNA ligase (NEB, UK) in a total reaction volume of 10µl. Control reaction was set up, to contain only digested vector with no shRNA plus ligase and used to check for successful ligations. The ligations were incubated overnight at 16°C. The following day, 5µl of each ligation reaction was transformed into a 50µl aliquot of MAX EFFICIENCY Stbl2 bacteria (Life Technologies, UK), as described in section 3.3.1.6. Individual colonies were picked and grown for precultures and subsequently minipreps and maxi preps were carried out for plasmid DNA extraction, as described in 3.3.1.8.

4.4.1.4. Screening for positive clones

Following transformation, colony isolation and plasmid purification by minipreps, samples were screened for the presence of an insert by restriction digest using the *XhoI* and *XbaI* restriction enzymes. The reaction consisted of 1x Buffer 4, 1x BSA, 1 unit of each enzyme and 18µl of miniprep DNA. Both enzyme restriction sites were present in the multiple cloning site of the vector and positives were analysed in comparison to empty vector. The expected sizes were as follows: For an empty vector: 334bp + 7.1kb; for vector + insert: 448bp + 7.1kb. Due to this very small difference, the digests were run on a 2% gel (w/v) for 2 hours at 80V. In addition, plasmids were sent for sequencing (Bioscience, UK) to verify each construct, using the primer that recognises the U6 promoter sequence in the vector:
5' – CGCACAGACTTGTGGGAGAAGCTCG – 3'.

4.4.2. Virus production

Lentiviruses were created using co-transfection (calcium phosphate co-precipitation) of shRNA MEF2A, shRNA scrambled and empty vector with the packaging plasmids into HEK 293T cells. HEK 293T cells are derived from human embryonic kidney cells (ATCC, USA) and are highly transfectable, because they express the SV-40 large T antigen, which allows for very high levels of plasmid replication. They were routinely grown in Falcon T75 flasks a complete media consisting of: DMEM containing 4.5g/L glucose and 2mM L-Glutamine supplemented with 10% heat inactivated FBS, 100U/ml Penicillin and 100µg/ml Streptomycin (Sigma, UK).

The day before transfection, $5x10^6$ cells were plated in 10ml of complete medium in 10cm dishes to reach 70-80% confluency. The following day, three hours prior to transfection, the medium was replaced with 5ml fresh culture medium. For each dish to be transfected, a separate reaction mix was prepared in a sterile 15ml Falcon tube, consisting of 10µg gene transfer plasmid, 5µg pRSV-Rev, 10µg pMDLG/pRRE and 3.5µg pMD2.G (packaging plasmids from Open Biosystems, Thermo Scientific, UK) topped up to 500 µl with sterile dH_2O. The above DNA assemble was mixed with 50µl 2.5mM $CaCl_2$ (Open Biosystems, UK). Whilst bubbling with a sterile, serological pipette, 500µl of 2x Hepes Buffer Saline (HBS) pH7.12 was added dropwise to the DNA/$CaCl_2$ mix. This was incubated at RT for 30 minutes. Shortly before transfection, 1µl of 100mM chloroquine was added to each culture dish that was to be transfected, followed by dropwise addition of 1ml of the DNA/$CaCl_2$/HBS complex and the cells were incubated overnight at 37°C, 5% CO_2. The following day, the medium was replaced with 10ml of fresh, complete medium and the plates were microscopically examined for the presence of GFP, an indicator of transfection efficiency. Supernatants containing lentiviral particles were collected 48 and 72 hours after transfection, pooled together, filtered through 0.22µm pore nitrocellulose filter and stored at -80°C. To concentrate the virus, the supernatant with lentiviral particles was transferred to ultracentrifugation tubes (Beckman Coulter) and spun down at 25000 RPM for 1.5h at 4^oC. The supernatant was carefully removed from the tubes without disturbing the lentiviral pellet. 30 µl of PBS was added to each tube and left overnight at 4^oC. The next day, the pellet was resuspended, aliquoted and stored at -80^o C.

4.4.3. Transduction of 293T cells and SPC-04 cells with lentiviruses silencing *MEF2A.*

To assess the effectiveness, specificity and knock-down efficiency of the shRNA constructs 293T cells were transduced with the lentivirus. The day prior infection, 293T cells were seeded in a 6-well plate at a concentration of 1 x 10^5 cells/well and incubated overnight. The following day, the media was changed to contain 2ml of lentiviral stock and 1ml of standard medium. To increase transduction efficiency, 3 µl of 8mg/ml polybrene (Sigma, USA) was added to each well. The plates were centrifuged for 90 minutes at 2500 RPM followed by 2.5h incubation at 37^o C, 5% CO_2, after which media was changed to 3ml of complete media and plates were returned to the incubator for a further 72h. Cells were then lysed for RNA extractions using the QIAGEN RNeasy Mini Kit (QIAGEN, UK), followed by cDNA synthesis (as described in 3.3.1.12 and 3.3.1.13).

Once transduction of 293T cells was confirmed and knock-down efficiency established, SPC-04 cells were infected with the concentrated virus, as described in 3.3.2.3. Briefly, the day before transduction, SPC-04 cells were seeded in 6-well laminin-coated Nunc plates at a density of 20000cells/cm^2. The following day, growth media was removed and replaced with RMM-containing 10µM DAPT and 100nM ATRA (pre-differentation stage) and a 1 µl of concentrated lentivirus per well. Cells containing lentiviral particles were incubated for 48h at 37°C, 5% CO_2, after which media was replaced with RMM- and the media was changed every 2 days until the end of the experiment. Differentiation day 1 was counted as the day after pre-differentiation had finished. Transduction efficiency was assessed using a fluorescent microscope (Leica DMIL) supplied with a Leica camera DFC420C (x10 objective).

RNA extracts were taken at the following time points: undifferentiated cells, pre-differentaited DAPT/ATRA stage, 7 days differentiation and 14 days differentiation. The cells were lysed for RNA extractions using the QIAgen RNeasy Mini Kit (QIAGEN, UK), followed by cDNA synthesis (as described in 3.3.1.12 and 3.3.1.13)

4.4.4. Quantification of *MEF2A* levels in 293T and SPC-04 cells.

cDNA samples were amplified by qPCR as described in 3.3.1.14. *GAPDH* was used as a house-keeping gene for extracts from 293T cells, and *RPL18* was used as a control for SPC-04 extracts. Sequences of all primers used are presented in a Table 4.4.

Table 4.4. Primer sequences used for qPCR amplification of *MEF2A* and the control gene *GAPDH* (for 293T cells) and *RPL18* for SPC-04 cells.

Primer		Sequence
GAPDH	Forward	5'-CATGAGAAGTATGACAACAGCCT -3'
	Reverse	5'-AGTCCTTCCACGATACCAAAGT-3'
RPL18	Forward	5' – GAGAGGTGTACCGGCATTTC – 3'
	Reverse	5' – CTCTGGCACGCTCGAACT – 3'
hMEF2A	Forward	5' – AGCCCTTCAAGGCTTCAACT – 3'
	Reverse	5' – GGTTCGGACTTGATGCTGAT – 3'

All qPCR data were analysed by a one-way ANOVA with *MEF2A* dosage and differentiation stage as factors and were followed post hoc analyses using Tukey's test. All statistical analyses were performed using IBM SPSS Statistics 20 software (IBM Corp., USA). A p-value <0.05 was considered statistically significant. The data was expressed as mean $\pm$ Standard Error of the Mean (SE).

4.4.5. RNA labelling and microarray hybridisation

The RNA, labelling, hybridisation, washing and scanning steps were outsourced and carried out by High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, Oxford. RNA integrity was analysed using Agilent RNA 6000 Nano Kit (Agilent Technologies, Inc; Germany) on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc). Total RNA was amplified and labelled (biotinylation) using the TargetAmp™-Nano Labeling Kit for Illumina Expression BeadChip (Cambio, UK) according to the manufacturer's instruction. Whole genome gene expression profiling was run on Illumina's HumanHT-12 v4.0 Expression BeadChip, which contains over 47000 probes. Samples were randomly allocated across the chip. The hybridised and washed chips were scanned using Illumina iScan Scanner.

4.4.6. Statistical and bioinformatic analyses of microarray data

Raw data was extracted using GenomeStudio Data Analysis Software, and was further processed by Dr. Sylvane Desrivières in R statistical environment (http://www.r-project.org)

using Lumi package. The variance stabilising transformation method was used, followed by quantile normalisation. Probes with unreliable expression measurements were flagged and removed from the analysis. Next, based on the coefficient of variation, genes which expression is constant across all the experiments (i.e., are invariant) were removed from the dataset. Of the 47,230 probes that feature on the Illumina platform, data from 10,069 probes were retained for further analysis. Using the Benjamini and Hochberg FDR method, q-values were generated in an attempt to correct for the effects of multiple testing, whereby expression differences producing q-values of $q < 0.05$ were considered true effects (Benjamini & Yekutieli, 2001). Transcripts from a total of 13 probes, representing 11 distinct genes were selected based on the above criteria.

Due to absence of a strong DE signal in this experiment, with only 11 genes meeting the criteria for $q < 0.05$, the uncorrected *p*-value threshold of $p < 0.005$ was chosen for subsequent functional analyses. This p-value is less stringent than FDR corrected $q < 0.05$ cutoffs and it has previously been used to identify differentially expressed genes in microarray studies (Miller, Woltjer, Goodenbour, Horvath, & Geschwind, 2013). Using this method, we generated a dataset containing 148 transcripts (representing 128 distinct genes) that were differentially expressed between *MEF2A* knock-down and non-silencing controls, which were subject to further functional analyses.

The functional annotation clustering tool, part of the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (http://david.abcc.ncifcrf.gov/), was used to identify enriched biological themes and functional-related gene groups in gene lists generated from microarray data (as described in 3.3.4). Additional analyses were performed using Ingenuity® systems software (http://www.ingenuity.com/) to identify relationships, mechanisms, functions, and pathways significantly enriched ($p<0.05$) in the dataset (as in 3.3.4). To identify if our set of differentially expressed (DE) genes were enriched for known susceptibility genes for autism or schizophrenia, the Simons Foundation Autism Research Initiative (SFARI) (https://sfari.org) and the Schizophrenia Gene Resource gene list were downloaded (SZ, http://www.szgene.org) were interrogated using the *phyper* function in R.

4.5. Results

4.5.1. *Mef2a* expression in zebrafish and mouse follows similar pattern to *Myt1l.*

MEF2A-MYT1L interaction during brain development in vertebrates was determined by performing qPCR analysis for Mef2a and compared its expression patterns to that of Myt1l in zebrafish and mouse.

Firstly, developing zebrafish embryos were screened for expression pattern of mef2aa gene. Mef2aa was chosen based on the literature as an ortholog of human MEF2A protein. Mef2ab was also identified bioinformatically in zebrafish, however it has no known biological function. Mef2a is located on chromosome 18 in the fish; it has 4 known transcripts, the longest spanning 2113bp.

qPCR time course analysis of the zebrafish *mef2a* expression pattern was performed with cDNA from 18SS, 24, 48 and 72h hpf-stage embryos. A high throughput analyses performed by Thisse *et al.* (B. Thisse, Thisse, C., 2004) showed that mef2aa starts being expressed between 20 somite stage (19hpf) and 5-Prim (24hpf), in various parts of the fish body, including the telencephalon. In agreement with this, we found that *mef2aa* expression was lowest at 18SS, increasing up to 2-fold at 24hpf and 48hpf and 6-fold at 72hpf (Figure 4.2). The results did not reach the significance level, when analysed by one-way ANOVA ($F_{(3,8)}$= 2.375, p=0.146). This could be due to only small changes in expression observed in these experiments. Nonetheless, the results confirm that mef2a is conserved throughout the fish development and its mRNA levels increase slightly with the embryo development. It is noteworthy that we only analysed extracts from the whole embryos, thus the observed changes in mRNA levels should not be merely attributed to differences observed within the brain, but the entire body.

From previous studies, we know that *mef2a* is highly expressed in the heart and somites during zebrafish embryogenesis (Y.-X. Wang et al., 2005), therefore increase in expression observed in our study could be partially attributed to that. However, it has also been shown that *mef2a* is expressed in brain during embryo development (Hammond & Udvadia, 2010). Looking at the ISH for *mef2a* as shown by Hammond *et al.* (Hammond & Udvadia, 2010) and comparing them to our *myt1l* ISH (see Chapter 2.4.1), we can observe similarities in brain expression patterns between *myt1l* and *mef2a*. Overlapping expression of Myt1l and Mef2a is most prominent in the telencephalon (Figure 4.3). Due to the widespread expression of *myt1l* in the brain at 48hpf, we can observe overlapping patterns in expression for both genes in

cerebellum, hindbrain and midbrain (Figure 4.3). The expression pattern of mef2a and myt1l the overlap in specific regions of CNS representing areas where mef2a could potentially regulate Myt1l activity in the developing brain.

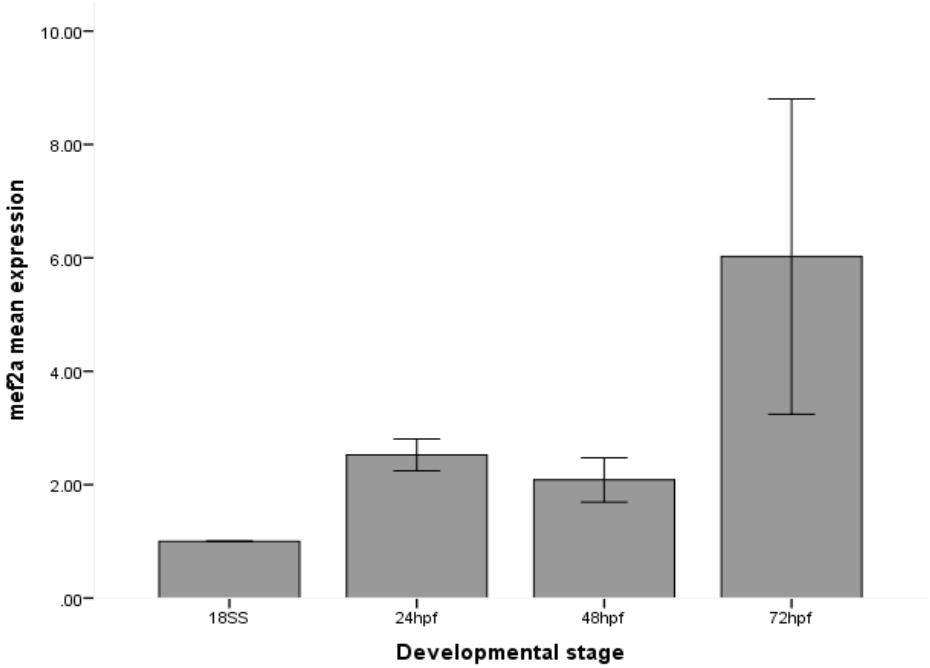


Figure 4.2. qPCR analysis of *mef2a* expression in zebrafish embryos. Highest expression (6-fold change) is reported from 3 day embryos. Results from 24hpf, 48hpf and 72hpf were calculated relative to the expression at 18SS stage. Data is presented as mean values of three independent experiments, error bars represent +/- 1SE of the mean.

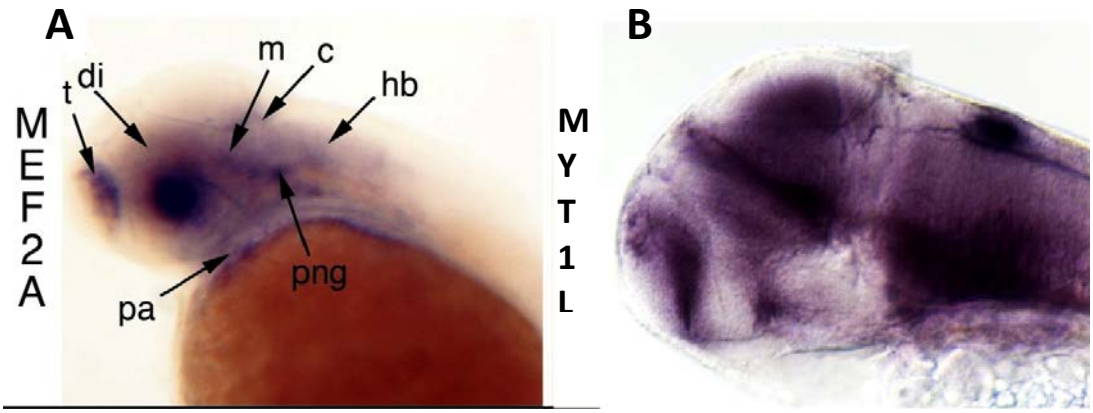


Figure 4.3. Mef2a and Myt1l expression overlaps in the central nervous system. At 48hpf, the overlap is most prominent in telencephalon (t), midbrain (m) and hindbrain (hb). **A** and **B** – lateral views of the head at 48hpf after whole mount ISH for Mef2a (taken from (Hammond & Udvadia, 2010)) **(A)** and myt1l (**B**). t-telenecephalon, di- diencephalon, m-midbrain, c-cerebellum, hb- hindbrain, pa- pharyngeal arches, png – peripheral nerve ganglia.

To validate these findings and to further assess changes of *Mef2a* transcript levels in the vertebrate brain, we investigated changes in expression of this transcription factor in the mouse brain during embryonic and postnatal development. The analysis showed that *Mef2a* expression is lower during embryonic development (E14), increasing just before the birth (E18) and reaching its highest levels 1 week after birth and remaining elevated at 1 month postnatally (Figure 4.4). A one-way ANOVA revealed a significant gene dosage effect in the mouse brain samples ($F_{(3,8)}$= 4.734, p=0.035). Post-hoc analysis using Tukey shown that the change was significant only between E10 and 1 week after birth (p=0.031), which corresponds to human third trimester (Romijn, Hofman, & Gramsbergen, 1991) when the brain growth spurt occurs (Clancy, Finlay, Darlington, & Anand, 2007). Inspection of the dissociation curves for zebrafish *mef2aa* and mouse *Mef2a* expression in RT-qPCR experiments revealed a single, specific amplicon for each measurement (Figure 4.5).

These findings are in general agreement with our results obtained while testing for *Myt1l* expression, when we observed significant increase in mRNA levels at stages E18 and 1 week postnatally (see Figure 2.10 in Chapter 2.4.3). To further elucidate the relationship between *Myt1l* and *Mef2a* expression, a Pearson correlation coefficient was computed. There was a strong positive correlation between the two variables (r=0.916, n=14, p=4.23E-06); increases in Myt1l expression were correlated with increases in Mef2a expression. This indicates that Mef2a follows similar expression pattern as Myt1l during the development of mouse brain.

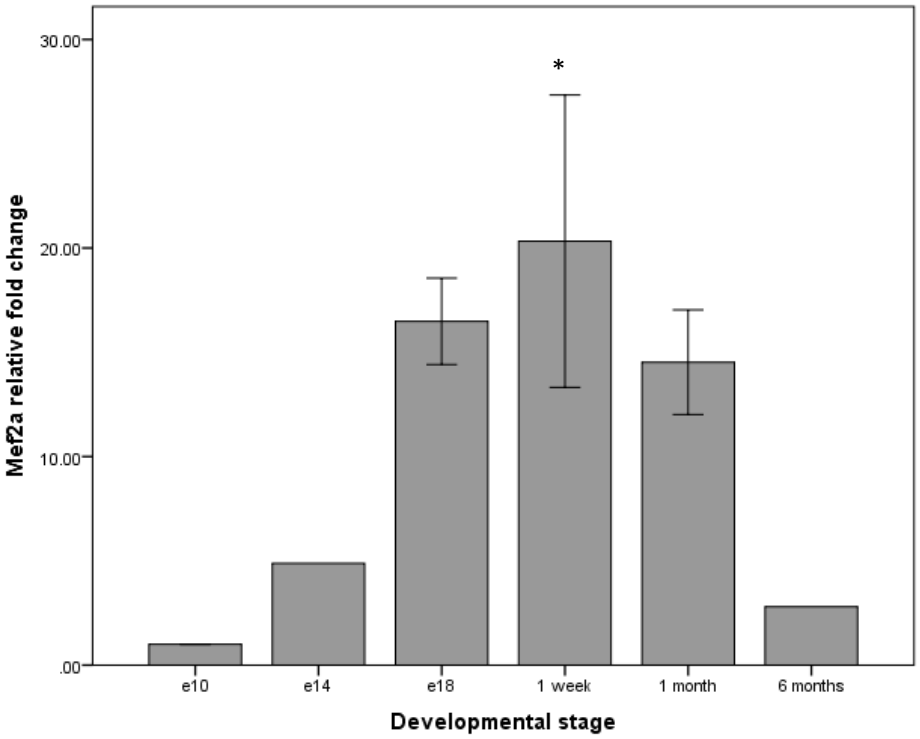


Figure 4.4. Graph representing stage specific expression of *Mef2a* during mouse brain development. The highest expression was recorded at post natal (P) week 1. Data from E14, E18, P1 week, P1 month and P6 months were calculated relative to the expression in the brain at E10. Statistical analysis (ANOVA) compared expression at E10, E18, 1 week and 1 month post natally. Tukey post hoc analyses showed significant increase in expression at 1 week after birth (*: p=0.035). Error bars represent +/- 1SE of the mean.

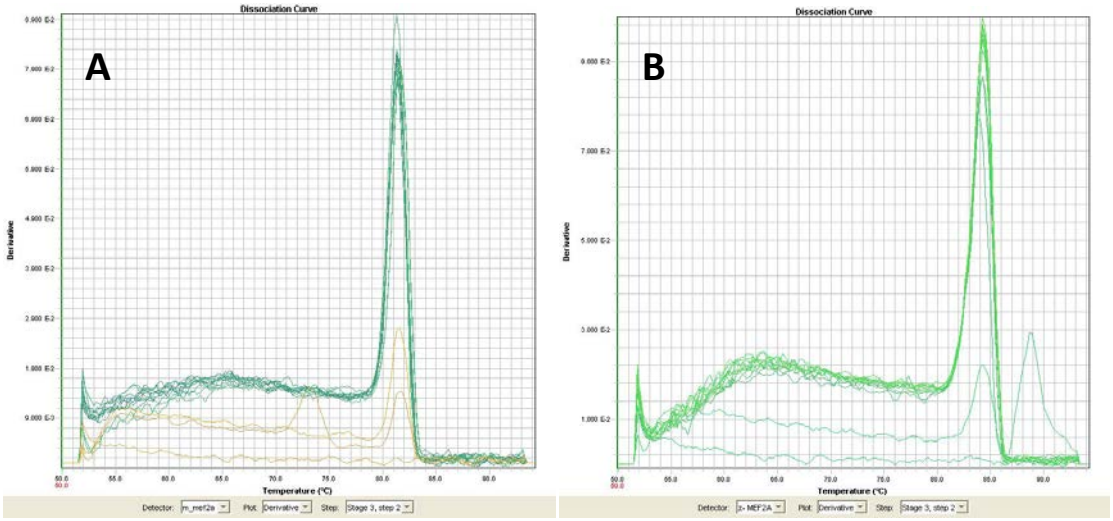


Figure 4.5. Confirmation of the specificity of *mef2a* RT-qPCR primers. Dissociation curves from qPCR analysis of Mef2a mRNA expression in zebrafish embryos (A) and in mouse brains (B) indicate a single, specific amplicon (i.e. one PCR product per primer pair). The additional peaks visible on both figures were generated from negative controls only (i.e. water).

4.5.2. Testing the efficiency of the *MEF2A* shRNA in HEK 293T cells

Before any attempts at knocking – down MEF2A were made, we needed to assess potential suitable cell lines that express *MEF2A* endogenously. Using qPCR, we have analysed SH-SY5Y, SK-N-SH, N2A and 293T cell lines for their endogenous expression of *MEF2A*. 293 T cells showed good expressions of MEF2A when their Ct values (Table 4.5) were compared to that of *GAPDH* (positive control). Although, the human neuroblastoma cell lines (SH-SY5Y and SK-N-SH) and the mouse neuroblastoma cells (N2A) were available in the lab and could have been suited for the shRNA validation purposes, *MEF2A* expression level was low in those cell lines compared to 293T cell line (Table 4.5). Thus, the latter was chosen to test MEF2A targeting shRNA for its efficiency.

Table 4.5. Ct values obtained from analysing *MEF2A* expression in four cell lines: SH-SY5Y, HEK 293T, SK-N-SH and N2A. Lower Ct values indicate higher levels of expression. The highest expression of *MEF2A* was observed in HEK 293T cells. The data were expressed as mean +/- 1SEM.

Cell line	Mean Ct *MEF2A*	Mean Ct *GAPDH*
SH-SY5Y	37.28 +/- 0.73	30.82 +/- 0.08
HEK 293T	29.86 +/- 0.14	17.06 +/- 0.46
SK-N-SH	Undetermined	16.36 +/- 0.15
N2A	32.18 +/- 0.06	28.01 +/- 0.03

293T cells were infected with lentiviruses containing shRNA *MEF2A* and shRNA scrambled as a control. We obtained on average a 30% decrease in *MEF2A* expression (Figure 4.6) when compared to the control. One-way ANOVA showed a significant knock-down of *MEF2A* by this shRNA ($F_{1,6}$=5.987, p=0.05).

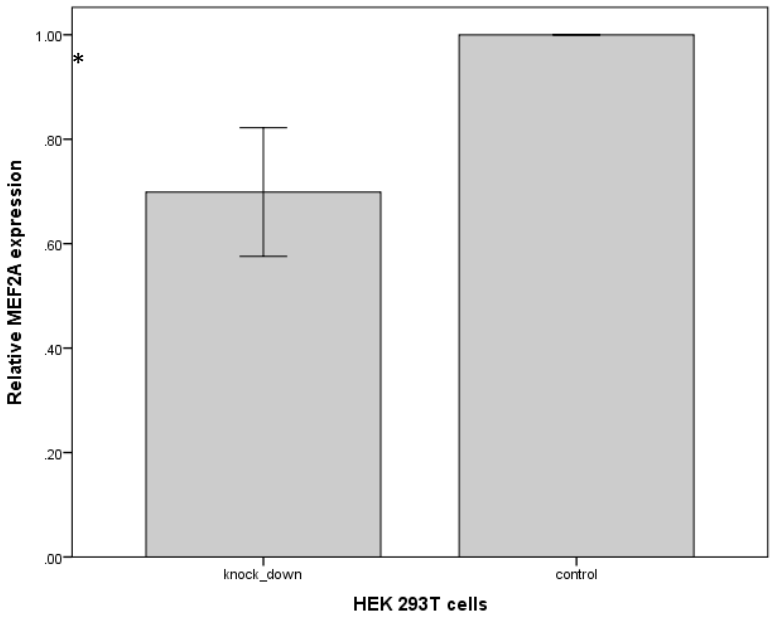


Figure 4.6. Relative fold expression of *MEF2A* mRNA in HEK 293T cells transduced with lentivirus silencing *MEF2A*. Cells were infected with virus carrying shRNA against MEF2A or scrambled shRNA, used as a control. Error bars represent +/- 1SE, *p=0.05.

4.5.3. Effects of *MEF2A* knock-down on SPC-04 cells

The human neural stem cell line SPC-04 was used as in vitro model for neural development. Firstly, we assessed endogenous *MEF2A* expression levels in SPC-04 at various stages of differentiation. By examining qPCR data of MEF2A expression during SPC-04 cells differentiation, we have found a good expression of this transcription factor across all the stages of stem cell differentiation (Figure 4.7), indicating endogenous expression of MEF2A mRNA in these cells. We detected a small decrease in expression at a pre-differentiation stage and 7 days of differentiation. A 1.5 fold increase at day 14 of differentiation compared to undifferentiated cells was observed. The small changes between the developmental stages were not significant when tested by one-way ANOVA ($F_{(4,14)}$=0.114, p=0.975).

Good endogenous gene expression across the developmental stages makes SPC-04 a good model to study gene silencing effects. We observed a lack of change in expression on mRNA level during the 14 day differentiation, which is in line with previous findings (S. W. Flavell et al., 2006; Steven W. Flavell et al., 2008; Scheschonka et al., 2007; Shalizi et al., 2006), which can suggest that unlike *MYT1L*, *MEF2A* might not be involved in the early stages of neural development but it might play a role later during synaptogenesis and post-synaptic neuronal differentiation, where MEF2 regulates synaptic number.

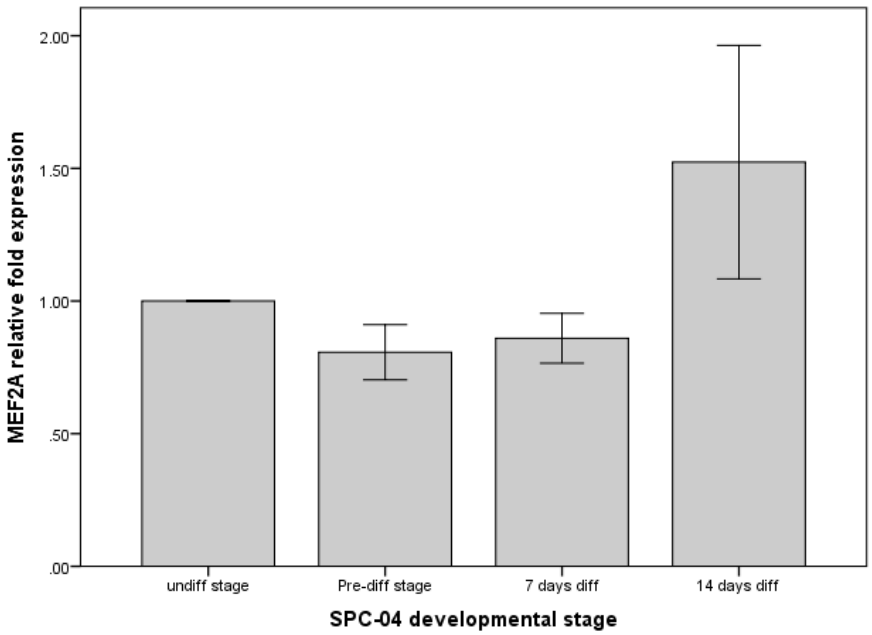


Figure 4.7. The relative fold expression of *MEF2A* in SPC-04 cells at different stages of differentiation. The data indicated that *MEF2A* is endogenously expressed at all tested stages of differentiation in neural progenitor/stem cells and the changes observed between different stages were not statistically significant (ANOVA, p=0.975). The error bars represent +/-1 SEM.

Having confirmed shRNA efficiency in 293T cells and endogenous expression of *MEF2A* in stem cells, SPC-04 cells were transduced with either *MEF2A* or non-silencing scrambled shRNAs and differentiated for 7 and 14 days. Microscopic examinations revealed a small change in cells morphology. Cells infected with shMEF2A (Figure 4.8 B,E) showed more developed neurites than the cells infected with a control shRNA (Figure 4.8 C,F). However, the phenotype is very subtle, and cells infected with MEF2A closely resemble uninfected, normally differentiating cells. Since this change is very subtle, it could be attributed to the effects of the virus alone, as we did not measure titers, therefore the virus containing control shRNA could have been more concentrated than virus carrying shRNA MEF2A.

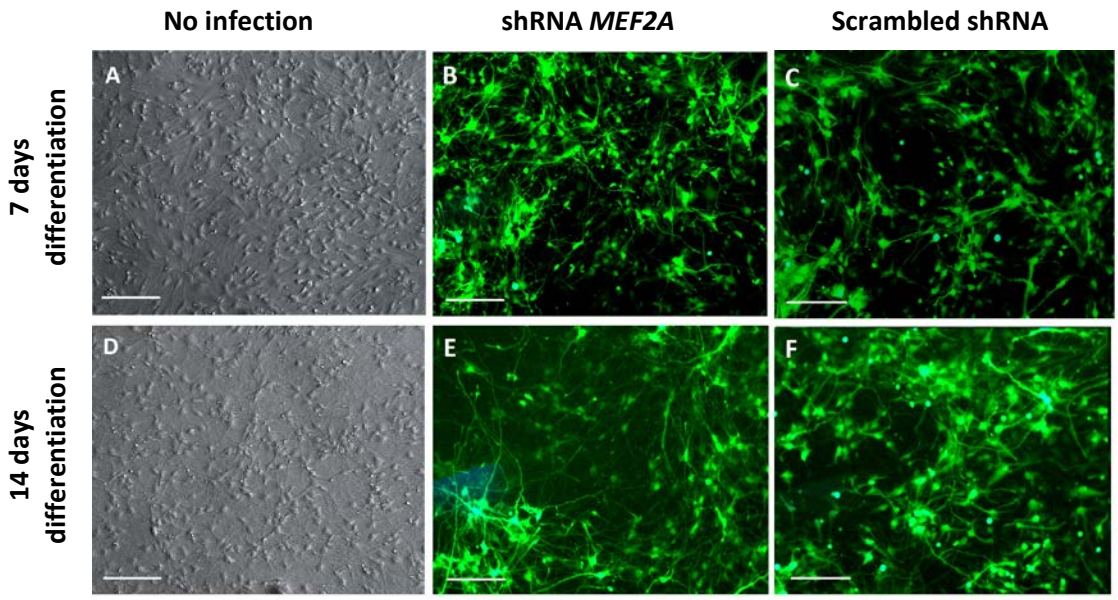


Figure 4.8. SPC-04 cells differentiation and transduction with lentivirus encoding shRNA *MEF2A* and scrambled shRNA. Each row represents a differentiation stage (7 or 14 days) and each column represents one condition obtained for this experiment. The first column contains phase-contrast pictures of normal differentiation occurring without transduction, second column represents cells treated with virus silencing MEF2A, and the last column corresponds to cells infected with scrambled shRNA. Figures 4.8. B-C and Figures 4.8 E-F are fluorescent microscope images indicating the GFP expression of transduced cells. Bar represents 200µm.

4.5.4. MEF2 gene product suppresses expression of *MYT1L*

The RNA harvested from transduced cells was first used to establish the knockdown efficiency of shMEF2A in SPC-04 cell line (3 samples at day 7 and 2 samples at day 14 of differentiation were analysed), which was on average a 40% reduction (Figure 4.9). The one-way ANOVA revealed a significant decrease ($F_{(2,14)}$=4.041, p=0.046) of *MEF2A* expression in all our samples, however due to only two replicates at 14 days of differentiation, we were unable to perform statistical analysis for that time point alone. Subsequent t-tests performed on the data collected from cells differentiated for 7 days revealed a significant change in MEF2A expression (t (3)= -3.279, p=0.023).

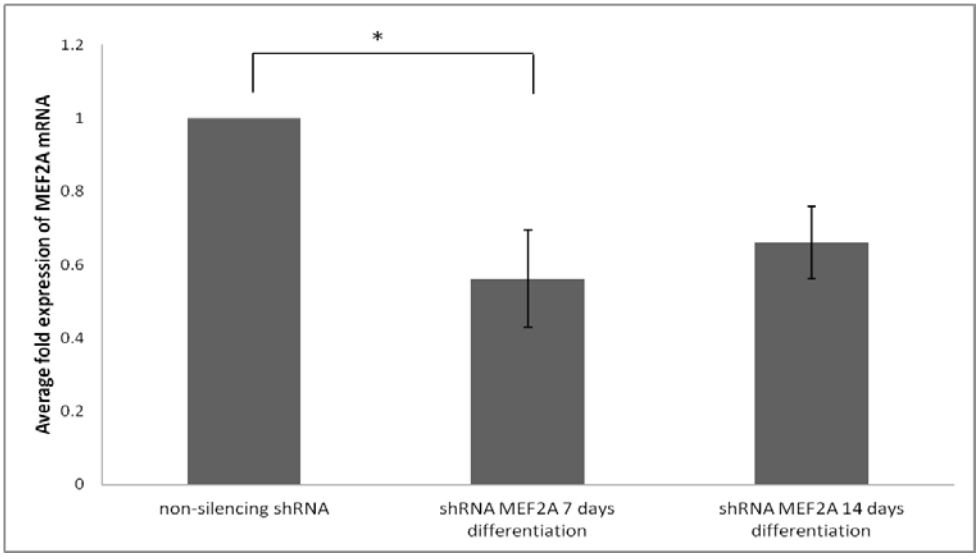


Figure 4.9. Relative fold expression of *MEF2A* mRNA in SPC-04 cells transduced with lentivirus silencing MEF2A for 7 and 14 days. Cells were infected with virus carrying shRNA against MEF2A or scrambled shRNA, used as a control. Error bars represent +/- 1SEM, *p<0.05

Investigation of potential relationship between *MYT1L* and *MEF2A* was obtained by measuring *MYT1L* mRNA levels in *MEF2A* knocked-down and *MEF2A* control samples and by comparing the expression of the two genes. Indeed, we found that association between the two variables was significant ($F_{(2,12)}$=7.029. p=0.01). Post-hoc Tukey revealed that *MYT1L* expression in *MEF2A* knock-down samples was significantly altered (p=0.013). In addition, the Pearson correlation coefficient was performed on data obtained from qPCR analysis of *MEF2A* knock-down and *MEF2A* control samples to assess the directionality of the effect that *MEF2A* has on MYT1L expression. The data showed strong negative correlation between the expression of the two genes in the *MEF2A* knock-down samples (r=-0.820, n=15, p=1.82E-04), suggesting that the higher the *MEF2A* knock-down efficiency (i.e. lower *MEF2A* expression) the higher *MYT1L* expression is. When the data on *MYT1L* and *MEF2A* expression from *MEF2A* knock-down experiments were plotted on the graph against each other, we found that *MYT1L* expression was dependant on *MEF2A* dosage.. For example, when *MEF2A* was silenced at 75%, *MYT1L* mRNA levels increased 27-fold, the trend followed the pattern that with lower MEF2A knock-down efficiency (i.e. higher expression on *MEF2A*) we observed lower MYT1L expression (Figure 4.10)

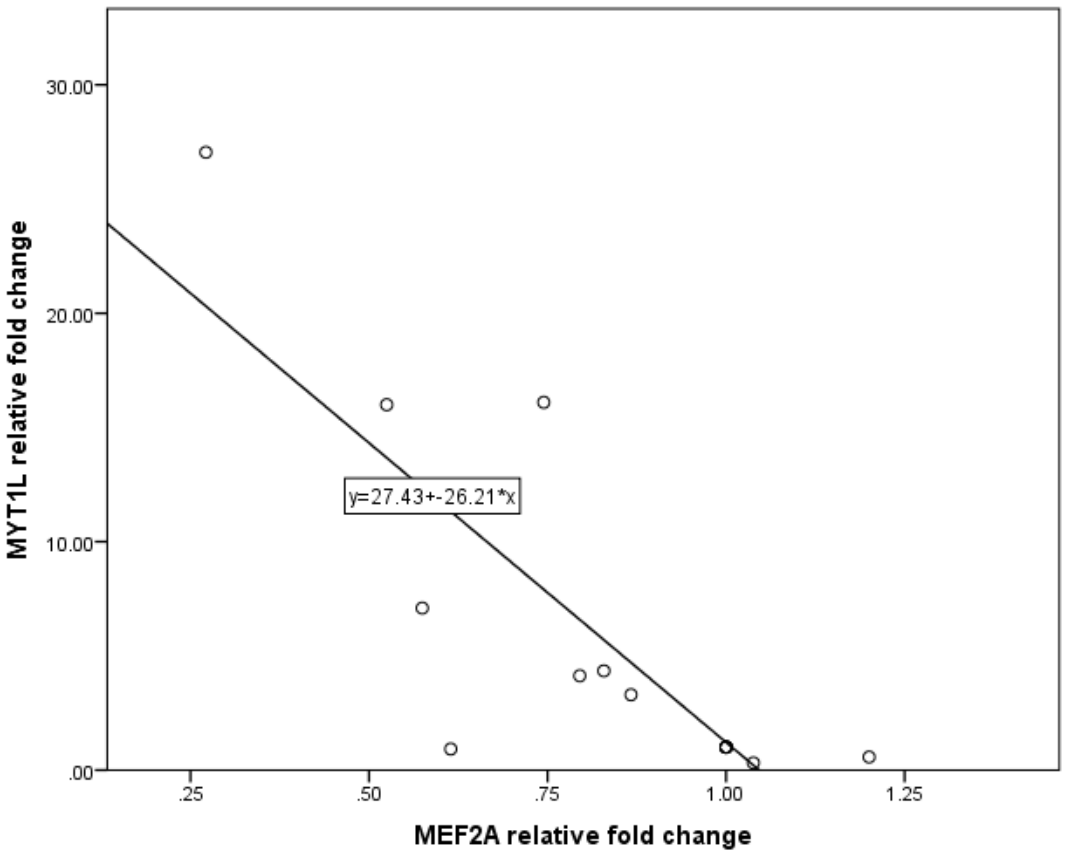


Figure 4.10. Graph representing *MYT1L* expression in *MEF2A* knock-down SPC-04 cells. *MEF2A* knock-down efficiency was assessed together with MYT1L mRNA transcript levels. The trend line represents the negative correlation between the two variables. MYT1L expression increased upon MEF2A silencing. We observed that with the increase of MEF2A knock-down efficiency (decrease of MEF2A expression), MYT1L mRNA levels raised.

We also checked if the relationship was mutual (i.e. if we observe altered MEF2A transcript levels dependent on *MYT1L* dosage), by measurement of the expression of *MEF2A* in the shMYT1L and MYT1L control samples. This analysis showed that in *MYT1L* knock-down samples, *MEF2A* levels remained unchanged ($F_{(1,12)}$=1.361, p=0.266). These findings suggest that *MEF2A* is a regulator of *MYT1L* by repressing the expression of MYT1L and thus potentially regulates the pathways by which *MYT1L* operate, which is a point of high interest in this research.

4.5.5. Differential gene expression analysis

In order to identify potential transcriptional targets of *MEF2A*, microarray gene expression profiling was performed using extracts from SPC-04 cells transduced with shRNA *MEF2A* or scrambled shRNA as a control and differentiated for 7 days. The data is representative of three independent experiments.

A total of 13 illumina probes detected significant ($q < 0.05$) differential mRNA expression between cells with silenced *MEF2A* and controls, which corresponded to transcripts from 11 distinct genes all of which were upregulated (table 4.6). Interestingly, seven of those genes (highlighted in yellow, table 4.6) had significant ($q<0.05$) differential expression in the *MYT1L* knock-down dataset. Worth noting, all the common genes that are present in both data sets, were down-regulated (chapter 3.4.3, Table 3.7) upon *MYT1L* silencing and up-regulated upon *MEF2A* reduction. This finding is in line with our hypothesis that *MEF2A* suppresses MYT1L, and by extension it is possible that it suppresses genes that are co-expressed with MYT1L.

Table 4.6. List of differentially expressed (FDR < 0.05) genes between SPC-04 cells with shRNA MEF2A and non-silencing controls. Relevant illumina probe, gene symbol and full name, log fold change, as well asuncorrected and corrected (FDR) *p*-values are listed for each gene. Genes are sorted according to *p*value; Highlighted in yellow are genes that were also DE in the MYT1L knock-down dataset.

Probe ID	Symbol	Definition	log fold change	p-value	FDR
6450097	FAM69C	chromosome 18 open reading frame 51 (C18orf51), mRNA.	1.79	6.92E-07	0.004589
2750176	AQP4	aquaporin 4 (AQP4), transcript variant a, mRNA.	1.80	1.76E-06	0.005907
1030021	SPARCL1	SPARC-like 1 (mast9, hevin) (SPARCL1), mRNA.	2.81	4.68E-06	0.011788
2100446	BCAN	brevican (BCAN), transcript variant 1, mRNA.	2.43	7.55E-06	0.015201
2760239	RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated) (RASGRP1), mRNA.	1.30	1.87E-05	0.0251
1110048	SCRG1	scrapie responsive protein 1 (SCRG1), mRNA.	1.38	2.25E-05	0.0251
520474	ODZ4	odz, odd Oz/ten-m homolog 4 (Drosophila) (ODZ4), mRNA.	1.85	2.39E-05	0.0251
1090326	TIMP4	TIMP metallopeptidase inhibitor 4 (TIMP4), mRNA.	1.25	2.49E-05	0.0251
3850059	LMO2	LIM domain only 2 (rhombotin-like 1) (LMO2), mRNA.	1.21	3.30E-05	0.030213
460544	CHODL	chondrolectin (CHODL), mRNA.	1.24	5.55E-05	0.04601
3460070	SPP1	secreted phosphoprotein 1 (SPP1), transcript variant 2, mRNA.	1.20	5.94E-05	0.04601

4.5.6. Functional annotation clustering and pathway analysis

The determination of whether the set of differentially expressed (DE) genes was enriched for any functional terms was performed by means of the functional annotation-clustering tool, which is part of the DAVID (v6.7) bioinformatic resource. As there is not a strong DE signal in this experiment, with only 11 significantly DE genes, the uncorrected p-value threshold of <0.005 was chosen as it is less stringent than both FDR corrected <0.05 cutoff. This way we generated a dataset containing 125 distinct genes, of which 107 were upregulated. 86 of those genes matched genes DE in *MYT1L* knock-down dataset (when dataset was widened to unadjusted p<0.005). All of the corresponding genes were upregulated in the *MEF2A* set but downregulated in *MYT1L*. These findings support our qPCR analysis suggesting that MEF2A may negatively regulate MYT1L.

Of the 147 probes submitted to DAVID, 125 were recognised as probing distinct genes in the human genome. A total of 56 functional gene clusters were generated from these set of DE genes, 5 of which reached the significant level of enrichment score (es) >1.3 ($p < 0.05$). Interestingly, the top scoring cluster, with ES=1.89 pertained to mainly neuron development and differentiation, axonogenesis and neuron projection (Figure 4.11). The enrichment score provides an indication of the biological significance of the gene groups being analysed. To better understand the details of enriched annotation terms associated with this gene list, functional annotation chart option in DAVID was implemented. When functional annotation chart option in DAVID was implemented it did not produce results that were significant after correction for a multiple testing. Briefly, the top ranked GO-term (GO:0030182)(Benjamini p= 0.298) was neuron differentiation which included 10 genes. Among the top scoring terms was also biological adhesion (GO:0022610) (p= 0.305) made up of 13 genes. This list of enriched GO terms largely overlaps with what was observed earlier with MYT1L knock-down dataset. It corroborates our hypothesis that MEF2A is a regulator of MYT1L and its co-expressed genes. It also suggests that MEF2A is involved in neural development and growth. The presence of cell adhesion molecules on the list together with neuron projection can imply that MEF2A plays a role in synaptic formation and the establishment of neuronal connectivity, which is what we would expect based on previous studies (S. W. Flavell et al., 2006; Lyons et al., 1995; Shalizi et al., 2006).

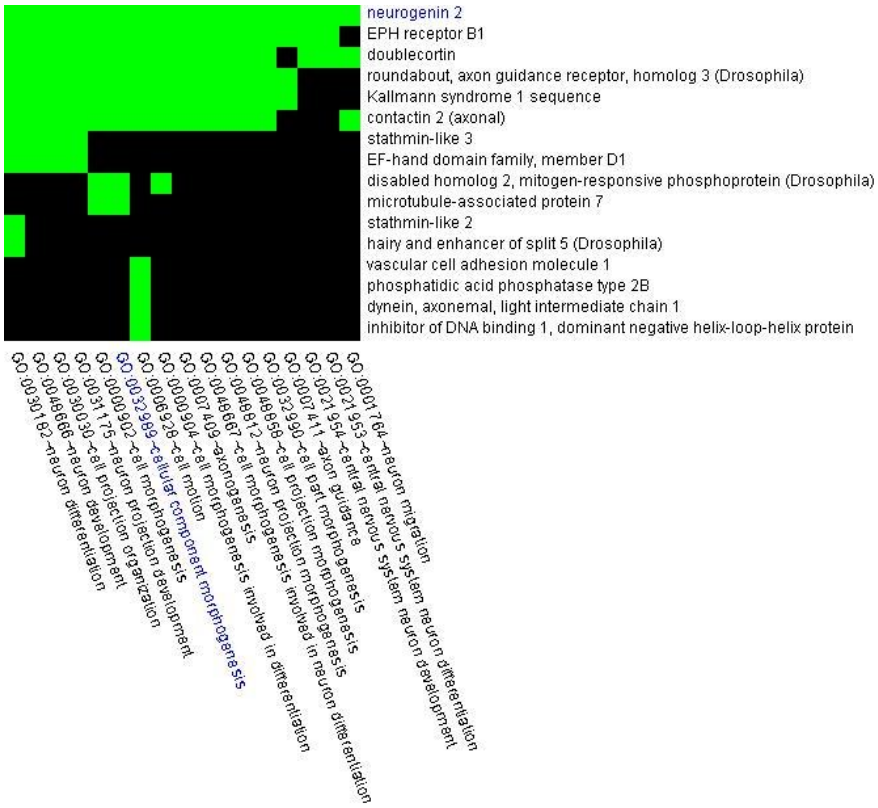


Figure 4.11. 2D view module from DAVID illustrating the top functional gene cluster pertaining mainly to neuron development, differentiation and neuron projection that was significantly enriched (ES>1.3) within the list of genes differentially expressed between SPC04 cells transduced with silencing *MEF2A* and non-silencing controls.

4.5.7. Ingenuity pathway analysis (IPA)

An analysis was made to determine whether this set of DE expressed genes was enriched in any biological functions, pathways or diseases Ingenuity pathway analysis (IPA) was performed. The genes that showed significant DE at an unadjusted p-value <0.005 were used for this analysis. Here, the top associated network was cardiovascular system development and function, cell to cell signalling and tissue development as 18 molecules in this network had altered expression in the knock-down cells (Figure 4.12). This is not entirely surprising, as *MEF2A* was primarily discovered and described in context of cardiac muscles development (Dodou, Sparrow, Mohun, & Treisman, 1995; Naya & Olson, 1999; Olson, 2003). Interestingly, the top subcategory within 'diseases and disorder' was Neurological Diseases, as 24 molecules previously associated with neurological disorders had altered expression in the *MEF2A* knock-down samples (Table 4.8). Based on our knowledge about MEF2A function, it is not surprising

to see that most of neurological disorders present on the list are of muscular nature (i.e. neuromuscular disease, movement disorders etc.). What is interesting is that neurodevelopmental disorders such as Schizophrenia and Rett Syndrome are also enriched with genes deregulated upon MEF2A silencing, suggesting that these disorders may be caused at least in part by disruption of MEF2A gene network.

Table 4.7 Ingenuity analyses showing the top diseases and disorders with significant ($p < 0.05$) enrichment of genes deregulated in *MEF2A* knock-down SPC-04 cells. The corresponding *p*-value and relevant gene symbols are also included.

Category	Diseases or Functions Annotation	p-Value	Molecules
Neurological Diseases	neuromuscular disease	1.83E-05	AEBP1,AGT,AQP4,BBOX1,DTNA,ETV5, GFAP,MAOB,PAQR6,SPARCL1,SPP1
	disorder of basal ganglia	5.23E-05	AEBP1,AQP4,BBOX1,DTNA,ETV5,GFAP, MAOB,PAQR6,SPARCL1,SPP1
	Movement Disorders	5.87E-05	AEBP1,AQP4,BBOX1,DTNA,ETV5,GFAP, MAOB,PAQR6,SPARCL1,SPP1
	progressive motor neuropathy	6.81E-05	AGT,AQP4,BBOX1,GFAP,PAQR6,SPARCL1, SPP1
	Parkinson's disease	4.53E-04	AQP4,BBOX1,PAQR6,SPARCL1,SPP1
	Huntington's Disease	8.01E-04	AEBP1,AQP4,BBOX1,DTNA,ETV5,GFAP, MAOB
	epilepsy	2.83E-03	DCX,KAL1
	Schizophrenia	1.09E-02	ACSBG1,GLRX,NELL2,PIK3R1,STMN2,SYT4
	Alzheimer's disease	1.14E-02	DCX,GABBR2,GFAP,GLRX,MAPK8IP1
	chronic fatigue syndrome	3.03E-02	BCAN,SPP1
	Rett Syndrome	4.45E-02	ATP1B1,GPR56,NELL2

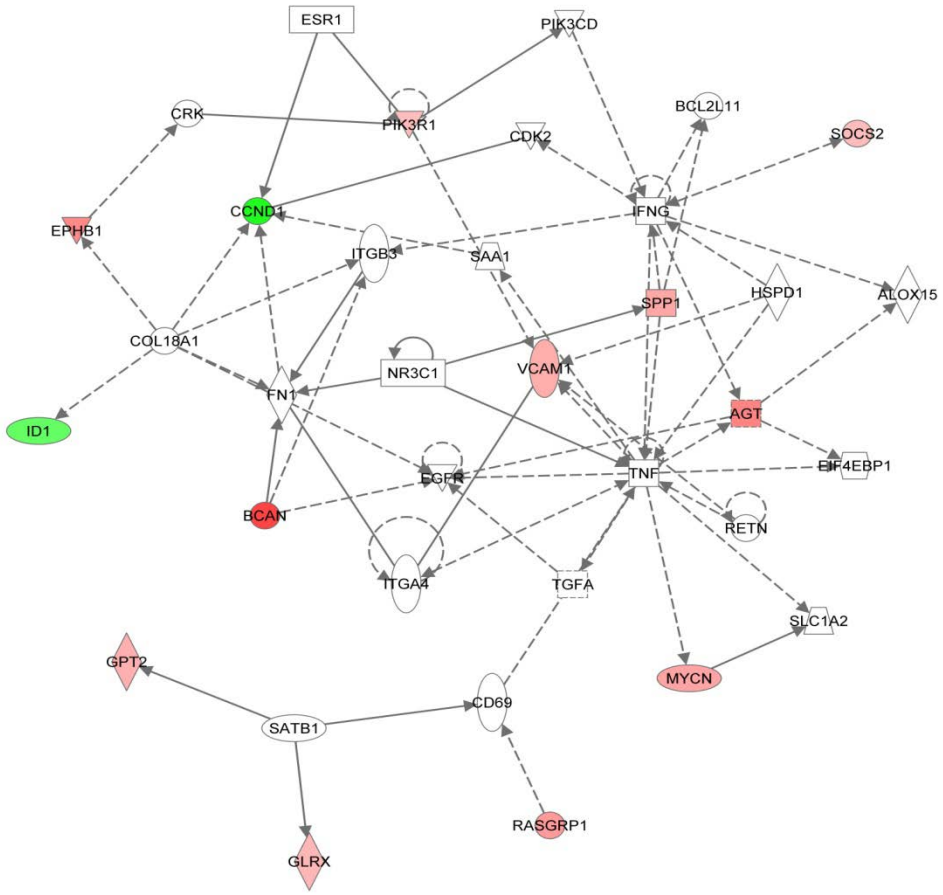


Figure 4.12. Top network in MEF2A knock-down SPC-04 cells representing cardiovascular system development and function, cell to cell signalling and tissue development. Green indicates genes which were significantly down-regulated in the MEF2A knock-down samples and red indicated genes which were significantly up-regulated in the MEF2A knock-down cells, compared to control cells.

4.5.8. Enrichment for schizophrenia (SZ) or autism (SFARI) genes

To further explore the potential relevance of deregulated genes in neurodevelopmental disorders, we analysed them for enrichment of genes that have previously been associated with ASD or schizophrenia (Table 4.9). The genes that showed significant DE at an unadjusted p-value <0.005 were used for this analysis. We have found 5 genes previously associated with ASD, 2 of them (*HEPACAM* and *NFIA*) were also DE in *MYT1L* knock-down dataset, however the enrichment wasn't significant. We found no genes enriched for schizophrenia.

Table 4.8. Enrichment of known ASD and SZ genes in the DE data set. The number of overlapping genes between the SFARI or SZ datasets and the DE genes list, and the p-values are included.

SFARI (no. of overlapping genes)	SFARI p-value	SZ (no. of overlapping genes)	SZ p-value
5 (*HEPACAM, DCX, MAOB, NFIA, ABAT)*	0.809	0	N/A

4.6. Discussion

In the nervous system, MEF2 transcription factors have emerged as regulators of activity-dependent neuronal survival and differentiation (S. W. Flavell et al., 2006; S. W. Flavell & Greenberg, 2008; Shalizi et al., 2006). The present series of experiments were designed to assess if MYT1L is a potential target gene of MEF2A during embryonic stem cells differentiation and development of zebrafish and mouse brain.

In this chapter, we found that *mef2a* expression in zebrafish embryos is similar to that of *myt1l*. These results only confirm mef2a presence in zebrafish embryos, and that it is evolutionary conserved gene. However based on our study only, we cannot comment on brain specific expression of this gene, as we tested whole embryos, and *mef2a* in contrast to *myt1l* is not a brain specific transcription factor. From the literature we know that, in the zebrafish head, *mef2a* is expressed at 48hpf in the telecephalon, midbrain and hindbrain, areas that were also highlighted by our ISH with *myt1l* at this time point. Thus we can suggest an overlap of the expression between the two genes, and possibility of their interaction in the aforementioned regions. However, data on mef2a expression in the zebrafish nervous system development is limited, and it would be indeed beneficial to further examine localisation of *mef2a* at various time points during zebrafish embryonic development, as it could provide us with valuable information of its brain region specific expression patterns and the potential similarities to *myt1l* expression. We have also found that in the developing mouse brain, Mef2a expression gradually increased during embryonic stages, however the peak in expression was observed at 1 week postnatally. Transcriptional levels of *Mef2a* remained elevated at 1 month and 6 months after birth, implicating that this transcription factor might have a role in synaptogenesis. This finding supports previous studies performed on rat brain lysates, where MEF2A and MEF2D proteins were identified during embryogenesis, but their expression increased during the first 3 weeks after birth, a time during which synapses form and are remodelled (S. W. Flavell et al., 2006). Zebrafish embryos develop to fully swimming larvae within 5 days pfs, and since we only tested up to 3 days pfs, I can only draw very broad conclusions on the similarities in Mef2a expression between the two species.

Results obtained from analysis of *mef2a* mRNA levels in the mouse brain are generally in agreement with the data obtained from fish embryos, as in both cases we observe gradual, modest increase in *Mef2a* expression, showing that the function of this gene is conserved in vertebrates. Additionally, when compared to *Myt1l*, *Mef2a* followed similar pattern of

expression. We observed the highest expression of *Myt1l* gene just prior to birth (>6000 fold increase), but its levels remained significantly elevated at 1 week and 1 month postnatally, suggesting that both of those genes can play a role during neuronal maturation and differentiation. It is noteworthy that *Myt1l* expression seems to be more abundant in both mouse brain and in the zebrafish than *Mef2a*. This could be due to the fact that *Myt1l* is brain specific transcription factor, while *Mef2a* has been found in other tissues (i.e. muscle tissues (Liu et al., 2012; Zhao, Zhao, & Peng, 2012). This can suggest that *Mef2a* has a more diverse role than *Myt1l* and that its primary function is to activate genes involved in muscle proliferation and differentiation.

We have also interrogated Gene Expression Nervous System Atlas (GENSAT) database (Gong et al., 2003) for possible presence of MEF2A or MYT1L however the expression of those two genes were not mapped in that database. The GENSAT project creates a library of bacterial artificial chromosome (BAC) clones that provide genetic access to each of the major cell populations in the mammalian brain, and it offers a collection of BAC transgenic mouse lines carrying fluorescent reporter genes that in turn allows further anatomical and physiological studies of these cells (Heintz, 2004). If Mef2a or Myt1l were present in the GENSET project it would help us to visualise and identify individual cell type expressing a gene of interest.

To elucidate if *MYT1L* may be a direct or indirect *MEF2A* target, we performed lentiviral mediated MEF2A knockdown in differentiating human neural embryonic stem cells, which resulted in an average of 40% reduction in *MEF2A* expression. Firstly, qPCR analysis in these samples revealed that decrease in *MEF2A* levels resulted in increased expression of *MYT1L*, whilst *MYT1L* knock-down did not affect *MEF2A* expression, which indeed suggests that *MYT1L* may be a direct or indirect *MEF2A* target. Microarray analysis revealed 11 differentially expressed genes between cells transduced with shRNA targeting *MEF2A* and non-silencing controls. The top hit was *FAM69C* gene, which is a member of the FAM69 family of cysteine-rich type II transmembrane proteins. It is found specifically in the brain and eye (Tennant-Eyles, Moffitt, Whitehouse, & Roberts, 2011), however its function remains yet to be determined.

Another gene that was deregulated upon *MEF2A* silencing was *AQP4 (*aquaporin 4). *AQP4* is abundant in the mammalian brain and it is the predominant water channel in the brain (Jérôme Badaut, Verbavatz, Freund-Mercier, & Lasbennes, 2000). Altered levels of AQP4 have been associated with acute brain injuries, such as trauma (Ke, Poon, Ng, Pang, & Chan, 2001), ischemia (De Castro Ribeiro, Hirt, Bogousslavsky, Regli, & Badaut, 2006) and subarachnoid

haemorrhage (J. Badaut et al., 2003). It has been reported that AQP4 expression was rapidly up-regulated in the astrocyte endfeet in the early stages of stroke onset. This increase in AQP4 was observed in the lesion site in a mouse stroke model, and the degree of its increase was temporally correlated with the degree of brain swelling (De Castro Ribeiro et al., 2006). Having AQP4 upregulated upon MEF2A silencing, can suggest that MEF2A may have a role in brain homeostasis. Interestingly, out of the 11 DE genes in MEF2A dataset, 7 were common to both datasets, *MEF2A* knock-and list and list of the DE genes generated from the *MYT1L* knock-down microarray experiment. This finding alone can imply that *MEF2A* and *MYT1L* operate through the same regulator pathways.

Gene set enrichment analyses were performed to ascertain whether differentially expressed genes were overrepresented in any particular biological functions, diseases, disorders, or canonical signalling pathways. The functional annotation analysis had shown that there are alterations in the expression of genes that have been previously associated with neuron projection, differentiation and development, indicating that MEF2A can play a role in neuronal differentiation through the regulation of those genes. Noteworthy, one of the genes repeatedly present in GO terms associated with neuron development and differentiation was NEUROD2. Neurogenic differentiation factor 2 (NeuroD2) is expressed exclusively in post-mitotic neurons and research indicates that it mediates neuronal differentiation (McCormick et al., 1996). From our study point of view, this is very interesting finding, as *NEUROD2*, in combination with *MYT1L* and *ASCL1*, was reported to facilitate micro-RNA mediated conversion of human fibroblasts into neurons (Yoo et al., 2011). The link between *NEUROD2, MYT1L* and *MEF2A* strengthens our hypothesis that *MYT1L* is a target of *MEF2A*.

Ingenuity Pathway analysis revealed that our set of deregulated genes was enriched for motor disorders. This finding is not surprising as MEF2 has been implicated in multiple aspects of muscle development (Potthoff & Olson, 2007). Our findings support previous studies on MEF2A, where MEF2A and MEF2D double knock-out mice showed deficits in motor coordination and enhanced short-term synaptic plasticity (Akhtar et al., 2012), whereas MEF2A and MEF2C were reported to be dysregulated in patients with myotonic dystrophy (Bachinski et al., 2010). We have also found genes that were previously associated with autism spectrum disorder, however most probably due to the sample size, the genes weren't significantly enriched. Nevertheless, all these results taken together suggest a list of potential target genes of MEF2A and they bring the interesting possibility of *MEF2A* being a regulator of *MYT1L*. We established that the majority of DE genes in the MEF2A dataset overlap with DE genes in the

MYT1L dataset, we also observed alterations in similar pathways or biological functions when the two sets were analysed. Furthermore, our results show that knock-down of MEF2A leads to increased expression of 107 genes and downregulation of 18 genes, which suggest that in a basal state those genes are maintained in a repressive state. Importantly, 87 of the genes upregulated upon MEF2A knockdown, were down-regulated by MYT1L silencing, but none was upregulated by the same treatment. These findings also suggest that upon reduction of *MEF2A* expression, genes belonging to the same regulatory pathway are activated to compensate for the loss of its function, whereas the same genes upon *MYT1L* silencing are also down-regulated, meaning that MYT1L can act like a regulatory 'switch' for those genes.

Previous studies have shown that MEF2s are required for neuronal survival and differentiation (Lyons et al., 1995; Mao et al., 1999). It has been shown that the reduction in MEF2s also prevents a metaplastic shift (i.e. activity-dependent regulation of neuronal plasticity), and reduction in MEF2 alone is sufficient to trigger a shift in plasticity responses (Chen et al., 2012). Thus, suggesting that MEF2 is a regulator of the plasticity threshold and implies orchestration of downstream effectors through altered transcriptional regulation. Here, by employing genome wide microarray analysis, we revealed target genes with diverse roles in neuron projection development and growth and neural differentiation.

There are some limitations to this study. The first limitation stems from microarray findings itself. We only found 11 genes that were significantly deregulated in the cells transduced with *MEF2A* shRNA compared to cells with scrambled shRNA. Similarly to *MYT1L*, we could not confirm results of reduced *MEF2A* in the cells transduced with shMEF2A compared to cells transduced with non-silencing shRNA. Thus, we were unable to confirm the efficiency of this transcription factor knock-down by microarray analysis. Although, we found by qPCR that *MEF2A* expression was altered in SPC-04 cells, our knock-down efficiencies were relatively low and fluctuating between the experiments, which could explain the microarray findings, and suggest little change between the two groups. One possible solution would have been to measure lentiviral concentrations and use a given concentration rather than certain quantity of the virus. This is an oversight on our part, as we performed experiments without the certain viral titer. The second limitation in this project is corroboration of microarray findings. Due to time constrains of these thesis, we were unable to validate any of our findings by qPCR. Similarly to MYT1L it would be interesting to assess the potential changes associated with

MEF2A down-regulation, it would be interesting to characterise neuronal properties of the cells with deregulated *MEF2A.*

In terms of future experiments, efforts should be made to construct a more reliable knock-down with a stable efficiency and known titer to corroborate our initial findings. Considering that increasing *MEF2* expression decreases the number of dendritic spines and excitatory synapses *in vitro* (S. W. Flavell et al., 2006), and blocks increases in spine density normally observed after repeated cocaine administration in rat nucleus accumbens neurons *in vivo* (Pulipparacharuvil et al., 2008), it would be interesting to study cell morphology and cellular properties of neurons past 7 days of differentiation. It remains unknown, how downregulation of *MEF2A* leads to increased levels of *MYT1L* and co-expressed genes. The mRNA experiments described in this chapter do not address whether *MYT1L* or the genes identified are direct or indirect targets of *MEF2A*. To determine which of these genes are directly regulated by MEF2A and to provide additional confirmation that our knock-down and microarray experiments have indeed identified MYT1L as a MEF2A target, chromatin immunoprecipitation (ChIP) could be performed. Also, our studies concentrated only on *in vitro* approach, therefore it would be important to extend our results to *in vivo* studies to further elucidate MEF2A function and to examine whether the candidate MEF2A target genes identified in these experiments are also deregulated upon MEF2A silencing in live neurons in the brain.

Dysregulation of transcription factor activity has been linked to various neuropsychiatric disorders, from substance abuse to neurodegenerative disorders (Steven W. Flavell et al., 2008; Eric M. Morrow et al., 2008; Pulipparacharuvil et al., 2008). It has been suggested, that MEF2A might be involved in such disorders, by for example regulating cocaine-induced changes in dendritic spines (Benjamini & Yekutieli, 2001), however the mechanisms that regulate MEF2A remain unknown. Determination of such mechanisms is required to unravel the roles of this transcription factor in neuronal and brain development in health and disease.

In conclusion, our study investigated the role of *MEF2A* on *MYT1L* expression during stem cell differentiation. It provided new insights into molecular interplay of these two transcriptional factors. To date, this is the first study that has showed a strong negative correlation between MEF2A and MYT1L mRNA expression levels. We have also identified sets of target genes that

are deregulated upon both *MEF2A* and *MYT1L* knockdown suggesting that they could be part of shared regulatory pathway.

5. Discussion

The general aim of this thesis was to investigate the functional role of a transcription factor *MYT1L.* This protein-coding gene is highly expressed in the brain, where it has been detected in neurons at the early stages of differentiation, but not in the glial lineage. This suggests that MYT1L has a role in the neurodevelopment (Armstrong et al., 1995; J. G. Kim et al., 1997). Additionally, studies have associated this gene with various neurodevelopmental and psychiatric disorders, such as autism (K. J. Meyer et al., 2012), schizophrenia (Y. Lee et al., 2012; Van Den Bossche et al., 2013; Vrijenhoek et al., 2008), intellectual disability (Bonaglia et al., 2009; S. J. Stevens et al., 2011) and major depressive disorder (T. Wang et al., 2010). The role of *MYT1L* in regulating neuron differentiation is further underlined by a recent breakthrough study showing that this protein, in conjunction with transcription factors A*scl1* and *Brn2*, is essential for inducing functional mature neurons in a mouse induced pluripotent stem cell model (T. Vierbuchen et al., 2010). However, the specific role of this gene and its mechanisms of action are not fully understood. To expand our knowledge about this transcription factor, the expression patterns were assessed in zebrafish and mouse brains, followed by gene silencing *in vitro* using differentiating stem cells as a model for neurodevelopment.

5.1. Role of myt1l in vertebrate brain development

In order to examine Myt1l expression patterns during embryonic development in vertebrates, zebrafish and mouse brains were used as a model.

The mechanisms underlying CNS development and specification of neuronal identity are well conserved across all vertebrates (Hauptmann & Gerster, 2000; Wilson & Houart, 2004) and development of the zebrafish is very similar to the embryogenesis in higher vertebrates, including humans. The short generation time, the large number of offspring and the transparency of the embryos make the zebrafish a good model organism to study the development using a genetic approach.

The patterns of expression of zebrafish *myt1l* were analysed using *in situ* hybridisation. *myt1l* expression was first detected at 24hpf in the telencephalon region. The telencephalon is involved in the zebrafish, as in other vertebrates, in high brain functions such as memory, basic emotions and precise motor control. In the developing fish, by 24hpf the first neuronal clusters

have extended axons and formed a simple scaffold of axon tracts and commissures connecting adjacent subdivisions. These first neurons are termed “primary neurons” and can be distinguished from later developing neurons. The first neurons of the telencephalon form shortly after 16hpf (Hjorth & Key, 2002), and they form a dorsorostal cluster and by 24hpf cells from this cluster extend their axons to create a new dorsoventrally directed tract (Ross, Parrett, & Easter, 1992).

As the fish development progressed, at 36, 48 and 72hpf, the expression of *myt1l* became more and more widespread and it was found in many regions of the brain, but it was restricted only to brain. The expression of *myt1l* during zebrafish development was also measured through quantitative RT-PCR experiments on pools of embryos at 18SS, 24, 48 and 72hpf. Myt1l mRNA levels remained low during the first 24hpf of development and then strongly increased between 24 and 72hpf. It is accepted that the fish hatches at 3days post fertilisation and it is called a larvae after that. During the hatching period the embryo continues to grow and morphogenesis of many organs is complete (Charles B. Kimmel et al., 1995). These results suggest that *myt1l* may play a pivotal role in brain development and that it might be involved in the process of neurogenesis. The data on zebrafish *myt1l* provided an insight into the embryonic expression profile of this gene.

In mouse and rat embryos Myt1l is solely transcribed in the developing central nervous system where it plays role in neurogenesis (Jiang et al., 1996; J. G. Kim et al., 1997; Weiner & Chun, 1997). In rodents, *Myt1l* is expressed predominantly in differentiating, postmitotic neurons, including those in the cerebral cortex, thalamus, hindbrain, and dorsal root ganglia (J. G. Kim et al., 1997). To investigate if Myt1l expression followed similar pattern in higher vertebrates, RNA extracts from mouse brains at various developmental time points were analysed. The analysis showed a very low expression at the early stages of development with the highest expression just before the birth. Although gradually decreasing, the mRNA levels remained elevated up to 6 months (last data point tested) after birth.

The mouse data corroborate findings obtained from the zebrafish model, suggesting that the most prominent role of Myt1l is just before birth, when neural development is most active with the peak of synapse formation and cortex development (Workman, Charvet, Clancy, Darlington, & Finlay, 2013). Unsurprisingly, *MYT1L* has been highly conserved during evolution with the human protein being 95% and 92% identical to mouse and rat *Myt1l* respectively (J. G. Kim & Hudson, 1992). Our findings support this hypothesis that its function is conserved in

vertebrates. These results also demonstrated that Myt1l is not only important for embryonic development, but it may also play role during post-natal brain development.

5.2. *MYT1L* gene targets

The pivotal role of *MYT1L* in regulating neuron differentiation is underlined by a recent study showing that the protein, in conjunction with *Brn2* and *Ascl1*, is sufficient for to directly reprogram embryonic and postnatal fibroblasts into functional neurons (T. Vierbuchen et al., 2010). *Ascl1* alone was able to produce cells with immature neuronal features, but co-infection with *Brn2* and *Myt1l* was required to produce cells with more mature neuronal features. Interestingly, these *Myt1l*-induced neurons resemble excitatory cortical neurons of the forebrain, mostly forming excitatory synapses (T. Vierbuchen et al., 2010). Moreover, *Myt1l*, in conjunction with *Ascl1, Brn2,* and *NeuroD1* (Z. P. Pang et al., 2011), or with *Ascl1*, *NeuroD2*, *miR-9/9** and *miR-124* (Yoo et al., 2011) can directly convert human fibroblasts into neural lineages. These induced human neurons show neuronal morphology and express multiple neuronal markers (Z. P. Pang et al., 2011; Pfisterer et al., 2011). They are able to generate action potentials and to receive synaptic contacts and can be directed towards distinct functional neurotransmitter phenotypes, such as dopaminergic neurons (Pfisterer et al., 2011) or glutamatergic neurons or GABAergic neurons (Z. P. Pang et al., 2011; Thomas Vierbuchen & Wernig, 2011; Yoo et al., 2011).

All of the aforementioned studies imply regulatory functions of *MYT1L* in neurogenesis, however, at present, there are no studies investigating the direct role of *MYT1L* in neurodevelopment. Many transcription factors are downstream targets of signalling pathways and integrate different signalling inputs that control cell behaviour (Armant et al., 2013). Although the concept of a key lineage-specifying transcription factors referred to as master regulators has been suggested (Oestreich & Weinmann, 2012; Yang et al., 2004), a growing body of research indicates that transcription factors act in a combinatorial fashion and the interplay between these factors determines the final outcome of the gene-expression profile (Davidson et al., 2002; Ravasi et al., 2010).

Our study is the first identifying possible target genes of *MYT1L*. Among the genes showing altered expression upon MYT1L silencing, the most significantly enriched GO term pertained to

genes belonging to cell adhesion and the top scoring cluster consisted of genes involved in synaptic transmission and regulation. Among the various proteins involved in the establishment of the neural networks, cell adhesion molecules play a pivotal role in the identification of the appropriate partner cell and the formation of a functional synapse (Brose, 1999). These findings suggest a role for *MYT1L* in processes such as neurite outgrowth, synaptic formation and neural migration. Further research is now required to verify and refine these initial findings.

Gene expression profiling revealed a list of potential targets of MYT1L, most of which were down-regulated upon MYT1L silencing, implying that the MYT1L regulates the expression of those genes. Perhaps underlying the potential role of MYT1L in neurodevelopment comes from close examination of the deregulated genes: *WNT7A* and *NLGN3*. WNT7A is a member of the Wnt family of signaling molecules and it has been shown to play an important role in axon development, guidance, and synapse formation (Lorenza Ciani et al., 2011; Gogolla, Galimberti, Deguchi, & Caroni, 2009; A. C. Hall, Lucas, & Salinas, 2000). Wnt7a induces axonal remodeling and synaptogenesis in cerebellar granule cells and in adult hippocampal neurons (Gogolla et al., 2009; A. C. Hall et al., 2000). Moreover, loss of Wnt7a expression led to reduced neural stem cell self-renewal, increased the rate of cell cycle exit in neural progenitors and reduced numbers of newborn neurons in the hippocampal dentate gyrus of adult mouse brains (Qu et al., 2013). In addition, the gene encoding a member of a family of neuronal cell-adhesion proteins, *NLGN3*, was downregulated upon MYT1L knock-down. Mutations in this gene have been identified in patients with ASDs (Jamain et al., 2003; Sudhof, 2008; Tabuchi et al., 2007). Autism is thought to arise from functional changes in neural circuitry and to be associated with an imbalance between excitatory and inhibitory synaptic transmission, however precise mechanism of action remain unknown (J. L. R. Rubenstein & Merzenich, 2003). Studies suggest that a defect in *NLGN3* may alter formation of specific synapses essential for the communication processes that are deficient in individuals with autistic spectrum disorder (Jamain et al., 2003). Recently it was found that induced neuronal (iN) cells that were derived from fibroblasts using Ascl1, Brn2 and Myt1l, when produced from a previously described mouse mutant in neuroligin-3, exhibited a phenotype similar to that observed in neurons with *Nlgn3* mutation (Chanda, Marro, Wernig, & Südhof, 2013). This study demonstrated that iN cells can be used for cellular disease modelling and analysing disease mechanisms.

5.3. Repressive function of *MEF2A* on *MYT1L*

MEF2A plays multiple roles in neuronal development, including neuronal survival (Mao et al., 1999), dendritic differentiation (Shalizi et al., 2006), synaptic density of hippocampal neurons (S. W. Flavell et al., 2006), spine density in nucleus accumbens (Pulipparacharuvil et al., 2008), and both synapse weakening and elimination and synaptic strengthening (Steven W. Flavell et al., 2008). A previous microarray study, conducted on differentiating stem cells (Desrivieres et al., 2014), revealed a set of genes that followed the same expression pattern as *MYT1L*. An examination of genes co-expressed with MYT1L during stem cells differentiation, led to the observation that the majority of them were enriched with MEF2A transcription factor binding sites. Based on that data and the hypothesis that MEF2A influences MYT1L expression, chapter 4 aimed at investigating regulatory role of *MEF2A* on *MYT1L*. One of the most intriguing questions raised by this study was whether MEF2A could act as a transcriptional repressor of MYT1L. To address this question the lentiviral-mediated *MEF2A* knockdown in differentiating human neuronal stem cells was used. Gene expression profiling experiments were performed on cells differentiated for 7 days. After successful validation of shRNA *MEF2A* knockdown efficiency, MYT1L mRNA levels were measured. Most interestingly, it was discovered that MYT1L levels were indeed dependant on MEF2A expression. A strong negative correlation was reported between the two genes, showing that MYT1L expression was enhanced upon MEF2A silencing. MYT1L levels responded increasingly to decreasing levels of MEF2A. It is very important to note that those correlations were observed only in *MEF2A* knock-down samples, as *MYT1L* silencing did not affect *MEF2A* expression, suggesting that *MYT1L* is a *MEF2A* target and not the other way around. Subsequent microarray analysis revealed a small number of differentially expressed genes in the *MEF2A* knock-down cells, all of which were upregulated. The DE analysis revealed that majority of genes deregulated upon knock-down of MEF2A were also de-regulated by knock-down of MYT1L, suggesting that both genes operate through similar or possibly the same regulatory pathways. Furthermore, the DE genes in this dataset were enriched for genes previously implicated in autism spectrum disorder, suggesting that MEF2A may be involved in key pathways related to these disorders.

Previous studies have shown that knockdown of MEF2A and MEF2D in cultured hippocampal neurons increases the number of excitatory synapses and the frequency of miniature excitatory postsynaptic currents (mEPSCs) (S. W. Flavell et al., 2006). In contrast, loss of

MEF2A in cerebellar granule neurons results in a decrease in the number of dendritic claws (Shalizi et al., 2006). Moreover, knock-out of the MEF2A-related transcription factor *MEF2C* in NSCs produced neurodevelopmental defects similar to ASD (Lipton et al., 2009). Additionally, mice with *MEF2C* conditionally knocked out at the NSC stage exhibited fewer, smaller, and more compacted neurons, similar to findings in Rett syndrome (Lipton et al., 2009). When *MEF2C* was knocked out later in development, neurogenesis was not affected, although synapse formation was altered (Lipton et al., 2009). Knockdown of MEF2A increases the density of orphan presynaptic sites in primary granule neurons and in the cerebellar cortex in rats (Yamada et al., 2013). Interestingly, the gene encoding synaptotagmin I (Syt1) has been recently found to be a novel direct repressed target gene of MEF2A in neurons (Yamada et al., 2013). It has been demonstrated that Syt1 repression mediates the ability of sumoylated MEF2A to eliminate orphan presynaptic sites (Yamada et al., 2013). These findings can possibly explain MYT1L relation to MEF2A. Findings from this thesis suggest another MEF2A target gene – *MYT1L*, where both *MEF2A* and *MYT1L* are components of a novel transcriptional pathway that orchestrates neuronal development.

5.4. Study limitations

Study limitations specific to each chapter has been discussed in the corresponding sections and they will not be recounted here.

One potential limitation to this thesis is that almost all of the work presented comes from *in vitro* studies. It would have been beneficial to assess the effects of *MYT1L* downregulation *in vivo.* Culturing cells *in vitro* provides a method of studying cells in a controlled environment. The major advantages of this system include the added control. However *in vitro* systems cannot recreate the complexity of *in vivo* biology. Currently, the extensive interactions among cells and tissues cannot be completely duplicated in a non-animal model. Most cell systems are representing only one cell type (no cell–cell interactions), often monoclonal in origin. Culture conditions are not homeostatic (i.e. exchange of media and continuous depletion of nutrients). However, using *in vivo* method isn't without flaws. The models require relative large amounts of test substance. The use of inbred-strains does not reflect natural variances (Hartung & Daston, 2009). Additionally, some of the genetic perturbations can be lethal to an organism, however in tissue culture the effects are much easier manipulated. *In vitro* models are relatively inexpensive, allow a high number of replicates and basic methodologies are widely

established. Also, novel technologies are quickly emerging which include image technologies as well as the diverse range of "omic" technologies. Taken together, the combination of *in vitro* and *in vivo* methods and findings can provide more evidence for gene function.

Another limitation to this thesis comes from qPCR experiments performed in the present study.*MYT1L* and *MEF2A* ΔCt values were normalised against the expression levels of only one housekeeping gene. Data normalisation in real-time q-PCR is a crucial step in gene quantification analysis. The appropriate choice of the internal controls (e.g. housekeeping genes) is critical for a meaningful quantitative RNA analysis. The most important characteristics of housekeeping genes are that they are present in all cells and that their expression levels remain relatively constant in different experimental conditions (Janssens, Janicot, Perera, & Bakker, 2004). However, in the present studies, stable expression of a control gene (such as *GAPDH* or *RPL18*) was only presumed and it was not verified experimentally. There is a growing body of evidence suggesting that at least two or three housekeeping genes should be used as internal standards because the use of a single gene for normalisation could lead to relatively large errors, as the expression of control genes may vary depending on the tissue types, experimental conditions and chosen time points (Vandesompele et al., 2002). GeNorm is commonly used software that utilises an algorithm for selecting the best candidate reference gene for a given experimental design (Vandesompele et al., 2002). The geometric mean of multiple carefully selected housekeeping genes has shown a vast improvement in the normalisation procedure (Vandesompele et al., 2002). Thus, the use of a single control gene in our experiments may not have been an optimal normalisation procedure, therefore future research with multiple housekeeping genes selected using the GeNorm algorithm to validate these findings would prove valuable.

5.5. Future directions

Due to lack of success in the identification of interacting partners for human MYT1L protein, future research should continue with this line of work. Most biological processes involve the action and regulation of multi-protein complexes. Protein-protein interactions regulate the activities of cells, therefore identifying them is critical to understanding cellular processes. Mass spectrometry techniques have been developed for large-scale screening to identify interacting proteins. Findings obtain by studying the protein-protein interactions could help

with the identification of interacting partners for MYT1L and should enable the formulation of hypotheses concerning MYT1L protein function and stimulate further research.

In chapter 2, we characterised myt1l expression during embryonic development of zebrafish. We used *in situ* hybridisation to localise *myt1l*. However, due to overstaining we were not able to determine exact expression patterns of myt1l gene product past 24hpf. One way to overcome this would be to look at the sections of brain tissue. To determine the specific brain regions and times at which *myt1l* is active in those regions would provide us with valuable clues of its possible role in neural development. Also, our findings from analysis of RNA from mouse brain suggest Myt1l involvement during post-natal development. In the zebrafish, we have only examined time points referred to embryonic development. It would be of interest to investigate the gene expression patterns in the fully developed zebrafish to assess myt1l levels and its distribution in adult zebrafish.

Additionally, it would be of interest to perform immunocytochemistry to investigate the consequences of MYT1L/MEF2A down-regulation on the SPC-04 cells differentiation potential. This technique would allow us to confirm expression of key neural stem cell marker. To achieve these cells would be stained with different markers alongside differentiation. Possible markers to use would be Nestin (neuronal stem cell marker), Glial fibrillary acidic protein (GFAP; marker of astrocytes), O1 (for oligodendrocytes), βIII-tubulin (for neurons) and Tau (mature neurons). These results would further expand our knowledge of the role of *MYT1L/MEF2A* during neuronal differentiation.

It remains unknown how *MEF2A* regulates expression of *MYT1L.* The mRNA profiling experiments described in these theses do not address the question whether the genes identified are direct or indirect targets of either MEF2A or MYT1L. To determine which of these genes are directly regulated by MEF2A/MYT1L and to provide additional confirmation that our microarray experiments have indeed identified MYT1L as a target of MEF2A, chromatin immunoprecipitation (ChIP) could be performed.

Lastly, our studies concentrated only on *in vitro* approach, therefore it would be important to extend our results to *in vivo* studies of further elucidate MYT1L role in brain development and to examine whether the candidate target genes identified in these experiments are also deregulated upon MYT1L or MEF2A silencing in live neurons in the brain. Manipulation of *MYT1L* expression in the brain will enable us to investigate the consequence of its absence in the developing brain and the possible phenotypes it might be involved in.

5.6. Conclusions

Despite the advances in science, molecular mechanisms underlying proper and abnormal brain remain largely unknown. Therefore, there is a need to identify genes and pathways involved in these processes. To summarise, using zebrafish as a model organism this thesis have shown when and where in brain MYT1L starts being expressed. Using lentiviral-mediated gene knockdown, effects of altered gene expression in differentiating stem cells were explored. As a result, this study provided foundations for dissecting role of MYT1L in neurodevelopment and its possible modes of action. This study has also supplemented the knowledge of the possible downstream targets of *MYT1L* during brain development. This is also the first study showing a direct link between MYT1L and MEF2A, demonstrating that MYT1L together with group of genes could be regulated via MEF2A. Therefore, the findings presented throughout this thesis contribute towards a better understanding of MYT1L role, function and its molecular mechanisms involved in the brain development.

References

Addington, A. M., & Rapoport, J. L. (2009). The genetics of childhood-onset schizophrenia: when madness strikes the prepubescent. *Curr Psychiatry Rep, 11*(2), 156-161.

Akhtar, M. W., Kim, M. S., Adachi, M., Morris, M. J., Qi, X., Richardson, J. A., . . . Monteggia, L. M. (2012). In vivo analysis of MEF2 transcription factors in synapse regulation and neuronal survival. *PLoS One, 7*(4), e34863. doi: 10.1371/journal.pone.0034863

Amatruda, J. F., & Patton, E. E. (2008) Chapter 1 Genetic Models of Cancer in Zebrafish. *Vol. 271* (pp. 1-34).

Amemori, T., Romanyuk, N., Jendelova, P., Herynek, V., Turnovcova, K., Prochazka, P., . . . Sykova, E. (2013). Human conditionally immortalized neural stem cells improve locomotor function after spinal cord injury in the rat. *Stem Cell Res Ther, 4*(3), 68. doi: 10.1186/scrt219

Andersen, S. L. (2003). Trajectories of brain development: point of vulnerability or window of opportunity? *Neuroscience & Biobehavioral Reviews, 27*(1–2), 3-18. doi: http://dx.doi.org/10.1016/S0149-7634(03)00005-8

Andrés, V., Cervera, M., & Mahdavi, V. (1995). Determination of the Consensus Binding Site for MEF2 Expressed in Muscle and Brain Reveals Tissue-specific Sequence Constraints. *Journal of Biological Chemistry, 270*(40), 23246-23249. doi: 10.1074/jbc.270.40.23246

Androutsellis-Theotokis, A., Leker, R. R., Soldner, F., Hoeppner, D. J., Ravin, R., Poser, S. W., . . . McKay, R. D. G. (2006). Notch signalling regulates stem cell numbers in vitro and in vivo. *Nature, 442*(7104), 823-826. doi: http://www.nature.com/nature/journal/v442/n7104/suppinfo/nature04940_S1.html

Armant, O., März, M., Schmidt, R., Ferg, M., Diotel, N., Ertzer, R., . . . Rastegar, S. (2013). Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos. *Dev Biol, 380*(2), 351-362. doi: http://dx.doi.org/10.1016/j.ydbio.2013.05.006

Armstrong, R. C., Kim, J. G., & Hudson, L. D. (1995). Expression of myelin transcription factor I (MyTI), a "zinc-finger" DNA-binding protein, in developing oligodendrocytes. *Glia, 14*(4), 303-321. doi: 10.1002/glia.440140407 [doi]

Armstrong, R. C., Migneault, A., Shegog, M. L., Kim, J. G., Hudson, L. D., & Hessler, R. B. (1997). High-grade human brain tumors exhibit increased expression of myelin transcription factor 1 (MYT1), a zinc finger DNA-binding protein. *J Neuropathol Exp Neurol, 56*(7), 772-781.

Avramopoulos, D. (2010). Genetics of psychiatric disorders methods: molecular approaches. *Psychiatr Clin North Am, 33*(1), 1-13. doi: 10.1016/j.psc.2009.12.006

Bachinski, L. L., Sirito, M., Böhme, M., Baggerly, K. A., Udd, B., & Krahe, R. (2010). Altered MEF2 isoforms in myotonic dystrophy and other neuromuscular disorders. *Muscle & Nerve, 42*(6), 856-863. doi: 10.1002/mus.21789

Back, S. A., & Plawner, L. L. (2012). Chapter 60 - Congenital Malformations of the Central Nervous System *Avery's Diseases of the Newborn (Ninth Edition)* (pp. 844-868). Philadelpia: W.B. Saunders.

Badaut, J., Brunet, J. F., Grollimund, L., Hamou, M. F., Magistretti, P. J., Villemure, J. G., & Regli, L. (2003). Aquaporin 1 and aquaporin 4 expression in human brain after subarachnoid hemorrhage and in peritumoral tissue. *Acta neurochirurgica. Supplement, 86*, 495-498.

Badaut, J., Verbavatz, J.-M., Freund-Mercier, M.-J., & Lasbennes, F. (2000). Presence of aquaporin-4 and muscarinic receptors in astrocytes and ependymal cells in rat brain: a clue to a common function? *Neuroscience Letters, 292*(2), 75-78. doi: http://dx.doi.org/10.1016/S0304-3940(00)01364-1

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medicine, 25*(01), 63-77. doi: doi:10.1017/S0033291700028099

Barinaga, M. (1998). Stroke-Damaged Neurons May Commit Cellular Suicide. *Science, 281*(5381), 1302-1303. doi: 10.1126/science.281.5381.1302

Barquinero, J., Eixarch, H., & Perez-Melgosa, M. (2004). Retroviral vectors: new applications for an old tool. *Gene Ther, 11 Suppl 1*, S3-9. doi: 10.1038/sj.gt.3302363

Bayer, S. A., Altman, J., Russo, R. J., & Zhang, X. (1993). Timetables of neurogenesis in the human brain based on experimentally determined patterns in the rat. *Neurotoxicology, 14*(1), 83-144.

Beattie, C. E., Carrel, T. L., & McWhorter, M. L. (2007). Fishing for a mechanism: Using zebrafish to understand spinal muscular atrophy. *Journal of Child Neurology, 22*(8), 995-1003.

Bell, C. C. (2002). Evolution of cerebellum-like structures. *Brain Behav Evol, 59*(5-6), 312-326.

Bell, C. C., Han, V., & Sawtell, N. B. (2008). Cerebellum-Like Structures and Their Implications for Cerebellar Function. *Annual Review of Neuroscience, 31*(1), 1-24. doi: doi:10.1146/annurev.neuro.30.051606.094225

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. 1165-1188. doi: 10.1214/aos/1013699998

Bergen, S. E., O'Dushlaine, C. T., Ripke, S., Lee, P. H., Ruderfer, D. M., Akterin, S., . . . Sullivan, P. F. (2012). Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry, 17*(9), 880-886. doi: 10.1038/mp.2012.73

Berkovits-Cymet, H. J., Amann, B. T., & Berg, J. M. (2004). Solution structure of a CCHHC domain of neural zinc finger factor-1 and its implications for DNA binding. *Biochemistry, 43*(4), 898-903. doi: 10.1021/bi035159d

Bier, E., & McGinnis, W. (2004). Model organisms in the study of development and disease.

Bill, B. R., Petzold, A. M., Clark, K. J., Schimmenti, L. A., & Ekker, S. C. (2009). A primer for morpholino use in zebrafish. *Zebrafish, 6*(1), 69-77. doi: 10.1089/zeb.2008.0555

Black, B. L., & Olson, E. N. (1998). Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu Rev Cell Dev Biol, 14*, 167-196. doi: 10.1146/annurev.cellbio.14.1.167

Blader, P., & Strähle, U. (2000). Zebrafish developmental genetics and central nervous system development. *Human Molecular Genetics, 9*(6), 945-951. doi: 10.1093/hmg/9.6.945

Blom, H. J., Shaw, G. M., den Heijer, M., & Finnell, R. H. (2006). Neural tube defects and folate: case far from closed. *Nat Rev Neurosci, 7*(9), 724-731. doi: http://www.nature.com/nrn/journal/v7/n9/suppinfo/nrn1986_S1.html

Bonaglia, M. C., Giorda, R., Massagli, A., Galluzzi, R., Ciccone, R., & Zuffardi, O. (2009). A familial inverted duplication/deletion of 2p25.1-25.3 provides new clues on the genesis of inverted duplications. *Eur J Hum Genet, 17*(2), 179-186. doi: 10.1038/ejhg.2008.160

Brose, N. (1999). Synaptic cell adhesion proteins and synaptogenesis in the mammalian central nervous system. *Naturwissenschaften, 86*(11), 516-524.

Burek, M., & Oppenheim, R. (1999). Cellular Interactions that Regulate Programmed Cell Death in the Developing Vertebrate Nervous System. In V. Koliatsos & R. Ratan (Eds.), *Cell Death and Diseases of the Nervous System* (pp. 145-179): Humana Press.

Cao, F., Xie, X., Gollan, T., Zhao, L., Narsinh, K., Lee, R. J., & Wu, J. C. (2010). Comparison of gene-transfer efficiency in human embryonic stem cells. *Mol Imaging Biol, 12*(1), 15-24. doi: 10.1007/s11307-009-0236-x

Caviness Jr, V. S., & Takahashi, T. (1995). Proliferative events in the cerebral ventricular zone. *Brain and Development, 17*(3), 159-163. doi: http://dx.doi.org/10.1016/0387-7604(95)00029-B

Chambers, S. M., Qi, Y., Mica, Y., Lee, G., Zhang, X. J., Niu, L., . . . Studer, L. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nat Biotechnol, 30*(7), 715-720. doi: 10.1038/nbt.2249

Chanda, S., Marro, S., Wernig, M., & Südhof, T. C. (2013). Neurons generated by direct conversion of fibroblasts reproduce synaptic phenotype caused by autism-associated neuroligin-3 mutation. *Proceedings of the National Academy of Sciences, 110*(41), 16622-16627. doi: 10.1073/pnas.1316240110

Chen, S. X., Cherry, A., Tari, P. K., Podgorski, K., Kwong, Y. K., & Haas, K. (2012). The transcription factor MEF2 directs developmental visually driven functional and structural metaplasticity. *Cell, 151*(1), 41-55. doi: 10.1016/j.cell.2012.08.028

Chico, T. J. A., Ingham, P. W., & Crossman, D. C. (2008). Modeling Cardiovascular Disease in the Zebrafish. *Trends in Cardiovascular Medicine, 18*(4), 150-155.

Chih, B., Afridi, S. K., Clark, L., & Scheiffele, P. (2004). Disorder-associated mutations lead to functional inactivation of neuroligins. *Human Molecular Genetics, 13*(14), 1471-1477. doi: 10.1093/hmg/ddh158

Ciani, L., Boyle, K. A., Dickins, E., Sahores, M., Anane, D., Lopes, D. M., . . . Salinas, P. C. (2011). Wnt7a signaling promotes dendritic spine growth and synaptic strength through Ca2+/Calmodulin-dependent protein kinase II. *Proceedings of the National Academy of Sciences, 108*(26), 10732-10737. doi: 10.1073/pnas.1018132108

Ciani, L., Boyle, K. A., Dickins, E., Sahores, M., Anane, D., Lopes, D. M., . . . Salinas, P. C. (2011). Wnt7a signaling promotes dendritic spine growth and synaptic strength through Ca(2)(+)/Calmodulin-dependent protein kinase II. *Proc Natl Acad Sci U S A, 108*(26), 10732-10737. doi: 10.1073/pnas.1018132108

Clancy, B., Finlay, B. L., Darlington, R. B., & Anand, K. J. (2007). Extrapolating brain development from experimental species to humans. *Neurotoxicology, 28*(5), 931-937. doi: 10.1016/j.neuro.2007.01.014

Cockrell, A., & Kafri, T. (2007). Gene delivery by lentivirus vectors. *Molecular Biotechnology, 36*(3), 184-204. doi: 10.1007/s12033-007-0010-8

Cocks, G., Romanyuk, N., Amemori, T., Jendelova, P., Forostyak, O., Jeffries, A. R., . . . Price, J. (2013). Conditionally immortalized stem cell lines from human spinal cord retain regional identity and generate functional V2a interneurons and motorneurons. *Stem Cell Res Ther, 4*(3), 69. doi: 10.1186/scrt220

Cohen Jr, M. M., & Sulik, K. K. (1992). Perspectives on holoprosencephaly: Part II. Central nervous system, craniofacial anatomy, syndrome commentary, diagnostic approach, and experimental studies. *Journal of Craniofacial Genetics and Developmental Biology, 12*(4), 196-244.

Collins, J. E., White, S., Searle, S. M., & Stemple, D. L. (2012). Incorporating RNA-seq data into the zebrafish Ensembl genebuild. *Genome Res, 22*(10), 2067-2078. doi: 10.1101/gr.137901.112

Conti, L., Pollard, S. M., Gorba, T., Reitano, E., Toselli, M., Biella, G., . . . Smith, A. (2005). Niche-Independent Symmetrical Self-Renewal of a Mammalian Tissue Stem Cell. *PLoS Biol, 3*(9), e283. doi: 10.1371/journal.pbio.0030283

Cowan, W. M., Fawcett, J. W., Oleary, D. D. M., & Stanfield, B. B. (1984). REGRESSIVE EVENTS IN NEUROGENESIS. *Science, 225*(4668), 1258-1265. doi: 10.1126/science.6474175

Craig, A. M., Graf, E. R., & Linhoff, M. W. (2006). How to build a central synapse: clues from cell culture. *Trends Neurosci, 29*(1), 8-20. doi: 10.1016/j.tins.2005.11.002

Crawford, T. Q., & Roelink, H. (2007). The Notch response inhibitor DAPT enhances neuronal differentiation in embryonic stem cell-derived embryoid bodies independently of sonic hedgehog signaling. *Developmental Dynamics, 236*(3), 886-892. doi: 10.1002/dvdy.21083

Cserjesi, P., & Olson, E. N. (1991). Myogenin induces the myocyte-specific enhancer binding factor MEF-2 independently of other muscle-specific gene products. *Mol Cell Biol, 11*(10), 4854-4862.

Czechowski, T., Bari, R. P., Stitt, M., Scheible, W.-R., & Udvardi, M. K. (2004). Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *The Plant Journal, 38*(2), 366-379. doi: 10.1111/j.1365-313X.2004.02051.x

Dalva, M. B., McClelland, A. C., & Kayser, M. S. (2007). Cell adhesion molecules: signalling functions at the synapse. *Nat Rev Neurosci, 8*(3), 206-220. doi: 10.1038/nrn2075

Daumke, O., Weyand, M., Chakrabarti, P. P., Vetter, I. R., & Wittinghofer, A. (2004). The GTPase-activating protein Rap1GAP uses a catalytic asparagine. *Nature, 429*(6988), 197-201. doi: http://www.nature.com/nature/journal/v429/n6988/suppinfo/nature02505_S1.html

Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C. H., . . . Bolouri, H. (2002). A genomic regulatory network for development. *Science, 295*(5560), 1669-1678. doi: 10.1126/science.1069883

De Castro Ribeiro, M., Hirt, L., Bogousslavsky, J., Regli, L., & Badaut, J. (2006). Time course of aquaporin expression after transient focal cerebral ischemia in mice. *Journal of Neuroscience Research, 83*(7), 1231-1240.

de Ligt, J., Willemsen, M. H., van Bon, B. W. M., Kleefstra, T., Yntema, H. G., Kroes, T., . . . Vissers, L. E. L. M. (2012). Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *New England Journal of Medicine, 367*(20), 1921-1929. doi: doi:10.1056/NEJMoa1206524

De Palma, M., & Naldini, L. (2002). [29] Transduction of a gene expression cassette using advanced generation lentiviral vectors. In M. I. Phillips (Ed.), *Methods in Enzymology* (Vol. Volume 346, pp. 514-529): Academic Press.

Desrivieres, S., Lourdusamy, A., Tao, C., Toro, R., Jia, T., Loth, E., . . . Schumann, G. (2014). Single nucleotide polymorphism in the neuroplastin locus associates with cortical thickness and intellectual ability in adolescents. *Mol Psychiatry*. doi: 10.1038/mp.2013.197

Dhara, S. K., & Stice, S. L. (2008). Neural differentiation of human embryonic stem cells. *J Cell Biochem, 105*(3), 633-640. doi: 10.1002/jcb.21891

Dietrich, J. B. (2013). The MEF2 family and the brain: from molecules to memory. *Cell Tissue Res, 352*(2), 179-190. doi: 10.1007/s00441-013-1565-2

Dimos, J. T., Rodolfa, K. T., Niakan, K. K., Weisenthal, L. M., Mitsumoto, H., Chung, W., . . . Eggan, K. (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science, 321*(5893), 1218-1221. doi: 10.1126/science.1158799

Dissen, G. A., Lomniczi, A., Neff, T. L., Hobbs, T. R., Kohama, S. G., Kroenke, C. D., . . . Ojeda, S. R. (2009). In vivo manipulation of gene expression in non-human primates using lentiviral vectors as delivery vehicles. *Methods, 49*(1), 70-77. doi: 10.1016/j.ymeth.2009.06.004

Dodou, E., Sparrow, D. B., Mohun, T., & Treisman, R. (1995). MEF2 proteins, including MEF2A, are expressed in both muscle and non-muscle cells. *Nucleic Acids Res, 23*(21), 4267-4274.

Doetsch, F., Caille, I., Lim, D. A., Garcia-Verdugo, J. M., & Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell, 97*(6), 703-716.

Dovey, H. F., John, V., Anderson, J. P., Chen, L. Z., De Saint Andrieu, P., Fang, L. Y., . . . Audia, J. E. (2001). Functional gamma-secretase inhibitors reduce beta-amyloid peptide levels in brain. *J Neurochem, 76*(1), 173-181. doi: 10.1046/j.1471-4159.2001.00012.x

Draghici, S., Khatri, P., Eklund, A. C., & Szallasi, Z. (2006). Reliability and reproducibility issues in DNA microarray measurements. *Trends in Genetics, 22*(2), 101-109. doi: http://dx.doi.org/10.1016/j.tig.2005.12.005

Driever, W., Solnica-Krezel, L., Schier, A. F., Neuhauss, S. C., Malicki, J., Stemple, D. L., . . . Boggs, C. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development, 123*, 37-46.

Duester, G. (2008). Retinoic acid synthesis and signaling during early organogenesis. *Cell, 134*(6), 921-931. doi: 10.1016/j.cell.2008.09.002

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D., & Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *Journal of Virology, 72*(11), 8463-8471.

Ertesvag, A., Naderi, S., & Blomhoff, H. K. (2009). Regulation of B cell proliferation and differentiation by retinoic acid. *Semin Immunol, 21*(1), 36-41. doi: 10.1016/j.smim.2008.06.005

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature, 292*(5819), 154-156.

Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A., & Sklar, P. (2005). Molecular Genetics of Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry, 57*(11), 1313-1323. doi: http://dx.doi.org/10.1016/j.biopsych.2004.11.024

Fisch, G. S. (2007). Animal models and human neuropsychiatric disorders. *Behav Genet, 37*(1), 1-10. doi: 10.1007/s10519-006-9117-0

Fishman, M. C. (1999). Zebrafish genetics: the enigma of arrival. *Proc Natl Acad Sci U S A, 96*(19), 10554-10556.

Flavell, S. W., Cowan, C. W., Kim, T. K., Greer, P. L., Lin, Y., Paradis, S., . . . Greenberg, M. E. (2006). Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number. *Science, 311*(5763), 1008-1012. doi: 10.1126/science.1122511

Flavell, S. W., & Greenberg, M. E. (2008). Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. *Annu Rev Neurosci, 31*, 563-590. doi: 10.1146/annurev.neuro.31.060407.125631

Flavell, S. W., Kim, T.-K., Gray, J. M., Harmin, D. A., Hemberg, M., Hong, E. J., . . . Greenberg, M. E. (2008). Genome-Wide Analysis of MEF2 Transcriptional Program Reveals Synaptic Target Genes and Neuronal Activity-Dependent Polyadenylation Site Selection. *Neuron, 60*(6), 1022-1038.

Földy, C., Malenka, Robert C., & Südhof, Thomas C. (2013). Autism-Associated Neuroligin-3 Mutations Commonly Disrupt Tonic Endocannabinoid Signaling. *Neuron, 78*(3), 498-509.

Follenzi, A., & Naldini, L. (2002). [26] Generation of HIV-1 derived lentiviral vectors. In M. I. Phillips (Ed.), *Methods in Enzymology* (Vol. Volume 346, pp. 454-465): Academic Press.

Fotos, J., Olson, R., & Kanekar, S. (2011). Embryology of the Brain and Molecular Genetics of Central Nervous System Malformation. *Seminars in Ultrasound, CT and MRI, 32*(3), 159-166. doi: http://dx.doi.org/10.1053/j.sult.2011.02.011

Gage, F. H. (2000). Mammalian neural stem cells. *Science, 287*(5457), 1433-1438.

Gécz, J., Shoubridge, C., & Corbett, M. (2009). The genetic landscape of intellectual disability arising from chromosome X. *Trends in Genetics, 25*(7), 308-316. doi: http://dx.doi.org/10.1016/j.tig.2009.05.002

Goff, S. P. (2001). Retroviridae: the retroviruses and their replication. In B. N. Fields, D. M. Knipe & P. M. Howley (Eds.), *Fields virology* (4th ed. / editors-in-chief, David M. Knipe, Peter M. Howley, associate editors, Diane E. Griffin ... [et al.] ed., pp. 1871–1939). Philadelphia ; London: Lippincott Williams & Wilkins.

Gogolla, N., Galimberti, I., Deguchi, Y., & Caroni, P. (2009). Wnt Signaling Mediates Experience-Related Regulation of Synapse Numbers and Mossy Fiber Connectivities in the Adult Hippocampus. *Neuron, 62*(4), 510-525.

Goldsmith, P. (2004). Zebrafish as a pharmacological tool: the how, why and when. *Curr Opin Pharmacol, 4*(5), 504-512. doi: 10.1016/j.coph.2004.04.005

Gong, S., Zheng, C., Doughty, M. L., Losos, K., Didkovsky, N., Schambra, U. B., . . . Heintz, N. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature, 425*(6961), 917-925. doi: http://www.nature.com/nature/journal/v425/n6961/suppinfo/nature02033_S1.html

Goridis, C., & Brunet, J.-F. (1999). Transcriptional control of neurotransmitter phenotype. *Current Opinion in Neurobiology, 9*(1), 47-53. doi: http://dx.doi.org/10.1016/S0959-4388(99)80006-3

Gotz, M., & Huttner, W. B. (2005). The cell biology of neurogenesis. *Nat Rev Mol Cell Biol, 6*(10), 777-788. doi: 10.1038/nrm1739

Graf, E. R., Zhang, X., Jin, S. X., Linhoff, M. W., & Craig, A. M. (2004). Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. *Cell, 119*(7), 1013-1026. doi: 10.1016/j.cell.2004.11.035

Greer, P. L., Hanayama, R., Bloodgood, B. L., Mardinly, A. R., Lipton, D. M., Flavell, S. W., . . . Greenberg, M. E. (2010). The Angelman Syndrome Protein Ube3A Regulates Synapse Development by Ubiquitinating Arc. *Cell, 140*(5), 704-716. doi: http://dx.doi.org/10.1016/j.cell.2010.01.026

Guerrini, R., Dobyns, W. B., & Barkovich, A. J. (2008). Abnormal development of the human cerebral cortex: genetics, functional consequences and treatment options. *Trends in Neurosciences, 31*(3), 154-162. doi: http://dx.doi.org/10.1016/j.tins.2007.12.004

Guerrini, R., & Parrini, E. (2010). Neuronal migration disorders. *Neurobiol Dis, 38*(2), 154-166. doi: http://dx.doi.org/10.1016/j.nbd.2009.02.008

Guo, S., Wilson, S. W., Cooke, S., Chitnis, A. B., Driever, W., & Rosenthal, A. (1999). Mutations in the Zebrafish Unmask Shared Regulatory Pathways Controlling the Development of Catecholaminergic Neurons. *Dev Biol, 208*(2), 473-487. doi: http://dx.doi.org/10.1006/dbio.1999.9204

Hacking, D. F. (2008). 'Knock, and it shall be opened': knocking out and knocking in to reveal mechanisms of disease and novel therapies. *Early Hum Dev, 84*(12), 821-827. doi: 10.1016/j.earlhumdev.2008.09.011

Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., Kane, D. A., . . . Nusslein-Volhard, C. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development, 123*, 1-36.

Hall, A. C., Lucas, F. R., & Salinas, P. C. (2000). Axonal Remodeling and Synaptic Differentiation in the Cerebellum Is Regulated by WNT-7a Signaling. *Cell, 100*(5), 525-535.

Hall, B. K. (2008). The neural crest and neural crest cells: discovery and significance for theories of embryonic organization. *J Biosci, 33*(5), 781-793.

Hammond, D. R., & Udvadia, A. J. (2010). Cabin1 expression suggests roles in neuronal development. *Developmental Dynamics, 239*(9), 2443-2451. doi: 10.1002/dvdy.22367

Hamshere, M. L., Walters, J. T., Smith, R., Richards, A. L., Green, E., Grozeva, D., . . . O'Donovan, M. C. (2013). Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry, 18*(6), 708-712. doi: 10.1038/mp.2012.67

Hartung, T., & Daston, G. (2009). Are In Vitro Tests Suitable for Regulatory Use? *Toxicological Sciences, 111*(2), 233-237. doi: 10.1093/toxsci/kfp149

Hauptmann, G., & Gerster, T. (2000). Regulatory gene expression patterns reveal transverse and longitudinal subdivisions of the embryonic zebrafish forebrain. *Mech Dev, 91*(1-2), 105-118.

He, X., & Rosenfeld, M. G. (1991). Mechanisms of complex transcriptional regulation: implications for brain development. *Neuron, 7*(2), 183-196.

Heidenreich, K., & Linseman, D. (2004). Myocyte enhancer factor-2 transcription factors in neuronal differentiation and survival. *Molecular Neurobiology, 29*(2), 155-165. doi: 10.1385/MN:29:2:155

Heintz, N. (2004). Gene Expression Nervous System Atlas (GENSAT). *Nat Neurosci, 7*(5), 483-483.

Hickey, M. A., & Chesselet, M.-F. (2003). Apoptosis in Huntington's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 27*(2), 255-265. doi: http://dx.doi.org/10.1016/S0278-5846(03)00021-6

Hjorth, J., & Key, B. (2002). Development of axon pathways in the zebrafish central nervous system. *Int J Dev Biol, 46*(4), 609-619.

Hong, E. J., West, A. E., & Greenberg, M. E. (2005). Transcriptional control of cognitive development. *Curr Opin Neurobiol, 15*(1), 21-28.

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., . . . Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature, advance online publication*. doi: 10.1038/nature12111

http://www.nature.com/nature/journal/vaop/ncurrent/abs/nature12111.html#supplementary-information

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols, 4*(1), 44-57. doi: http://www.nature.com/nprot/journal/v4/n1/suppinfo/nprot.2008.211_S1.html

Huang, J., Perlis, R. H., Lee, P. H., Rush, A. J., Fava, M., Sachs, G. S., . . . Smoller, J. W. Cross-disorder genomewide analysis of schizophrenia, bipolar disorder, and depression. *Am J Psychiatry, 167*(10), 1254-1263. doi: appi.ajp.2010.09091335 [pii]

10.1176/appi.ajp.2010.09091335 [doi]

Ito, M. (2002). Historical Review of the Significance of the Cerebellum and the Role of Purkinje Cells in Motor Learning. *Annals of the New York Academy of Sciences, 978*(1), 273-288. doi: 10.1111/j.1749-6632.2002.tb07574.x

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nat Rev Neurosci, 9*(4), 304-313.

Jamain, S., Quach, H., Betancur, C., Rastam, M., Colineaux, C., Gillberg, I. C., . . . Bourgeron, T. (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet, 34*(1), 27-29. doi: 10.1038/ng1136

Janssens, N., Janicot, M., Perera, T., & Bakker, A. (2004). Housekeeping genes as internal standards in cancer research. *Mol Diagn, 8*(2), 107-113.

Jiang, Y., Yu, V. C., Buchholz, F., O'Connell, S., Rhodes, S. J., Candeloro, C., . . . Rosenfeld, M. G. (1996). A novel family of Cys-Cys, His-Cys zinc finger transcription factors expressed in developing nervous system and pituitary gland. *J Biol Chem, 271*(18), 10723-10730.

Kano, M., & Hashimoto, K. (2009). Synapse elimination in the central nervous system. *Curr Opin Neurobiol, 19*(2), 154-161. doi: 10.1016/j.conb.2009.05.002

Kari, G., Rodeck, U., & Dicker, A. P. (2007). Zebrafish: An emerging model system for human disease and drug discovery. *Clinical Pharmacology and Therapeutics, 82*(1), 70-80.

Kaufman, M. H. (1992). *The Atlas of Mouse Development*: Academic Press.

Ke, C., Poon, W. S., Ng, H. K., Pang, J. C.-S., & Chan, Y. (2001). Heterogeneous responses of aquaporin-4 in oedema formation in a replicated severe traumatic brain injury model in rats. *Neuroscience Letters, 301*(1), 21-24. doi: http://dx.doi.org/10.1016/S0304-3940(01)01589-0

Kile, B. T., & Hilton, D. J. (2005). The art and design of genetic screens: mouse. *Nat Rev Genet, 6*(7), 557-567. doi: 10.1038/nrg1636

Kim, J. G., Armstrong, R. C., v Agoston, D., Robinsky, A., Wiese, C., Nagle, J., & Hudson, L. D. (1997). Myelin transcription factor 1 (Myt1) of the oligodendrocyte lineage, along with a closely related CCHC zinc finger, is expressed in developing neurons in the mammalian central nervous system. *J Neurosci Res, 50*(2), 272-290. doi: 10.1002/(SICI)1097-4547(19971015)50:2<272::AID-JNR16>3.0.CO;2-A [pii]

Kim, J. G., & Hudson, L. D. (1992). Novel member of the zinc finger superfamily: A C2-HC finger that recognizes a glia-specific gene. *Mol Cell Biol, 12*(12), 5632-5639.

Kim, S. U. (2004). Human neural stem cells genetically modified for brain repair in neurological disorders. *Neuropathology, 24*(3), 159-171. doi: 10.1111/j.1440-1789.2004.00552.x

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental Dynamics, 203*(3), 253-310. doi: 10.1002/aja.1002030302

Kimmel, C. B., Hatta, K., & Eisen, J. S. (1991). Genetic control of primary neuronal development in zebrafish. *Development, Suppl 2*, 47-57.

Kintner, C. (2002). Neurogenesis in embryos and in adult neural stem cells. *J Neurosci, 22*(3), 639-643.

Ko, M. S. H. (2001). Embryogenomics: developmental biology meets genomics. *Trends in Biotechnology, 19*(12), 511-518. doi: http://dx.doi.org/10.1016/S0167-7799(01)01806-6

Lanz, T. A., Guilmette, E., Gosink, M. M., Fischer, J. E., Fitzgerald, L. W., Stephenson, D. T., & Pletcher, M. T. (2013). Transcriptomic analysis of genetically defined autism candidate genes reveals common mechanisms of action. *Mol Autism, 4*(1), 45. doi: 10.1186/2040-2392-4-45

Lee, Y., Mattai, A., Long, R., Rapoport, J. L., Gogtay, N., & Addington, A. M. (2012). Microduplications disrupting the MYT1L gene (2p25.3) are associated with schizophrenia. *Psychiatric Genetics, 22*(4), 206-209 210.1097/YPG.1090b1013e328353ae328353d.

Lee, Y., Mattai, A., Long, R., Rapoport, J. L., Gogtay, N., & Addington, A. M. (2012). Microduplications disrupting the MYT1L gene (2p25.3) are associated with schizophrenia. *Psychiatr Genet, 22*(4), 206-209. doi: 10.1097/YPG.0b013e328353ae3d

Lichtman, J. W., & Colman, H. (2000). Synapse elimination and indelible memory. *Neuron, 25*(2), 269-278.

Lipton, S. A., Li, H., Zaremba, J. D., McKercher, S. R., Cui, J., Kang, Y. J., . . . Nakanishi, N. (2009). Autistic phenotype from MEF2C knockout cells. *Science, 323*(5911), 208. doi: 10.1126/science.323.5911.208b

Littlewood, T. D., Hancock, D. C., Danielian, P. S., Parker, M. G., & Evan, G. I. (1995). A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Research, 23*(10), 1686-1690.

Liu, Y., Niu, W., Wu, Z., Su, X., Chen, Q., Lu, L., & Jin, W. (2012). Variants in exon 11 of MEF2A gene and coronary artery disease: evidence from a case-control study, systematic review, and meta-analysis. *PLoS One, 7*(2), e31406. doi: 10.1371/journal.pone.0031406

Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Eun, A., . . . Peng, J. (2003). 15000 unique zebrafish EST clusters and their future use in microarray for profiling gene expression patterns during embryogenesis. *Genome Res, 13*(3), 455-466. doi: 10.1101/gr.885403

Lyons, G. E., Micales, B. K., Schwarz, J., Martin, J. F., & Olson, E. N. (1995). Expression of mef2 genes in the mouse central nervous system suggests a role in neuronal maturation. *J Neurosci, 15*(8), 5727-5738.

Mao, Z., Bonni, A., Xia, F., Nadal-Vicens, M., & Greenberg, M. E. (1999). Neuronal activity-dependent cell survival mediated by transcription factor MEF2. *Science, 286*(5440), 785-790.

Mao, Z., & Wiedmann, M. (1999). Calcineurin enhances MEF2 DNA binding activity in calcium-dependent survival of cerebellar granule neurons. *J Biol Chem, 274*(43), 31102-31107.

Matsunami, N., Hadley, D., Hensel, C. H., Christensen, G. B., Kim, C., Frackelton, E., . . . Hakonarson, H. (2013). Identification of rare recurrent copy number variants in high-risk autism families and their prevalence in a large ASD population. *PLoS One, 8*(1), e52239. doi: 10.1371/journal.pone.0052239

Matthews, K., Christmas, D., Swan, J., & Sorrell, E. (2005). Animal models of depression: navigating through the clinical fog. *Neurosci Biobehav Rev, 29*(4-5), 503-513. doi: 10.1016/j.neubiorev.2005.03.005

McCarroll, S. A., Murphy, C. T., Zou, S., Pletcher, S. D., Chin, C. S., Jan, Y. N., . . . Li, H. (2004). Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat Genet, 36*(2), 197-204. doi: 10.1038/ng1291

McCormick, M. B., Tamimi, R. M., Snider, L., Asakura, A., Bergstrom, D., & Tapscott, S. J. (1996). NeuroD2 and neuroD3: distinct expression patterns and transcriptional activation potentials within the neuroD gene family. *Mol Cell Biol, 16*(10), 5792-5800.

McKay, R. (1997). Stem Cells in the Central Nervous System. *Science, 276*(5309), 66-71. doi: 10.1126/science.276.5309.66

McKinsey, T. A., Zhang, C. L., & Olson, E. N. (2002). MEF2: a calcium-dependent regulator of cell division, differentiation and death. *Trends Biochem Sci, 27*(1), 40-47.

McTigue, D. M., & Tripathi, R. B. (2008). The life, death, and replacement of oligodendrocytes in the adult CNS. *J Neurochem, 107*(1), 1-19. doi: 10.1111/j.1471-4159.2008.05570.x

Meyer, K. J., Axelsen, M. S., Sheffield, V. C., Patil, S. R., & Wassink, T. H. (2012). Germline mosaic transmission of a novel duplication of PXDN and MYT1L to two male half-siblings with autism. *Psychiatr Genet, 22*(3), 137-140. doi: 10.1097/YPG.0b013e32834dc3f5

Meyer, K. J., Axelsen, M. S., Sheffield, V. C., Patil, S. R., & Wassink, T. H. (2012). Germline mosaic transmission of a novel duplication of PXDN and MYT1L to two male half-siblings with autism. *Psychiatric Genetics, 22*(3), 137-140 110.1097/YPG.1090b1013e32834dc32833f32835.

Miller, J. A., Woltjer, R. L., Goodenbour, J. M., Horvath, S., & Geschwind, D. H. (2013). Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med, 5*(5), 48. doi: 10.1186/gm452

Mitalipov, S., & Wolf, D. (2009). Totipotency, pluripotency and nuclear reprogramming. *Adv Biochem Eng Biotechnol, 114*, 185-199. doi: 10.1007/10_2008_45

Miyata, T., Kawaguchi, A., Okano, H., & Ogawa, M. (2001). Asymmetric Inheritance of Radial Glial Fibers by Cortical Neurons. *Neuron, 31*(5), 727-741. doi: http://dx.doi.org/10.1016/S0896-6273(01)00420-2

Morris, J. A. (2009). Chapter 11 - Zebrafish: a model system to examine the neurodevelopmental basis of schizophrenia. In A. Sawa (Ed.), *Progress in Brain Research* (Vol. Volume 179, pp. 97-106): Elsevier.

Morrow, E. M., Yoo, S.-Y., Flavell, S. W., Kim, T.-K., Lin, Y., Hill, R. S., . . . Walsh, C. A. (2008). Identifying Autism Loci and Genes by Tracing Recent Shared Ancestry. *Science, 321*(5886), 218-223. doi: 10.1126/science.1157657

Morrow, E. M., Yoo, S. Y., Flavell, S. W., Kim, T. K., Lin, Y., Hill, R. S., . . . Walsh, C. A. (2008). Identifying autism loci and genes by tracing recent shared ancestry. *Science, 321*(5886), 218-223. doi: 10.1126/science.1157657

Mothe, A. J., & Tator, C. H. (2012). Advances in stem cell therapy for spinal cord injury. *J Clin Invest, 122*(11), 3824-3834. doi: 10.1172/JCI64124

Mouse Genome Sequencing, C., Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., . . . Lander, E. S. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature, 420*(6915), 520-562. doi: 10.1038/nature01262

Mueller, T., & Wullimann, M. F. (2003). Anatomy of neurogenesis in the early zebrafish brain. *Developmental Brain Research, 140*(1), 137-155. doi: http://dx.doi.org/10.1016/S0165-3806(02)00583-7

Nadarajah, B. (2003). Radial glia and somal translocation of radial neurons in the developing cerebral cortex. *Glia, 43*(1), 33-36. doi: 10.1002/glia.10245

Nadarajah, B., & Parnavelas, J. G. (2002). Modes of neuronal migration in the developing cerebral cortex. *Nat Rev Neurosci, 3*(6), 423-432.

Nasevicius, A., & Ekker, S. C. (2000). Effective targeted gene 'knockdown' in zebrafish. *Nat Genet, 26*(2), 216-220. doi: 10.1038/79951

Naya, F. J., & Olson, E. (1999). MEF2: a transcriptional target for signaling pathways controlling skeletal muscle growth and differentiation. *Curr Opin Cell Biol, 11*(6), 683-688.

Neale, B. M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K. E., Sabo, A., . . . Daly, M. J. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature, 485*(7397), 242-245. doi: 10.1038/nature11011

Nelson, H. C. (1995). Structure and function of DNA-binding proteins. *Curr Opin Genet Dev, 5*(2), 180-189.

Newman, M., Musgrave, F. I., & Lardelli, M. (2007). Alzheimer disease: Amyloidogenesis, the presenilins and animal models. *Biochimica et Biophysica Acta - Molecular Basis of Disease, 1772*(3), 285-297.

Nielsen, J. A., Berndt, J. A., Hudson, L. D., & Armstrong, R. C. (2004). Myelin transcription factor 1 (Myt1) modulates the proliferation and differentiation of oligodendrocyte lineage cells. *Mol Cell Neurosci, 25*(1), 111-123. doi: 10.1016/j.mcn.2003.10.001 [doi]

S1044743103003191 [pii]

Novara, F., Beri, S., Giorda, R., Ortibus, E., Nageshappa, S., Darra, F., . . . Van Esch, H. (2010). Refining the phenotype associated with MEF2C haploinsufficiency. *Clin Genet, 78*(5), 471-477. doi: 10.1111/j.1399-0004.2010.01413.x

Nüsslein-Volhard, C. (1995). The identification of genes controlling development in flies and fishes. *Les Prix Nobel*(273).

Oestreich, K. J., & Weinmann, A. S. (2012). Master regulators or lineage-specifying? Changing views on CD4+ T cell transcription factors. *Nat Rev Immunol, 12*(11), 799-804. doi: 10.1038/nri3321

Olson, E. N. (2003). Coronary artery disease and the MEF2A transcription factor. *Sci Aging Knowledge Environ, 2003*(48), pe33. doi: 10.1126/sageke.2003.48.pe33

Oppenheim, R. W. (1991). Cell Death During Development of the Nervous System. *Annual Review of Neuroscience, 14*(1), 453-501. doi: doi:10.1146/annurev.ne.14.030191.002321

Pang, Z. P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D. R., Yang, T. Q., . . . Wernig, M. (2011). Induction of human neuronal cells by defined transcription factors. *Nature, 476*(7359), 220-223. doi: http://www.nature.com/nature/journal/v476/n7359/abs/nature10202.html#supplementary-information

Pang, Z. P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D. R., Yang, T. Q., . . . Wernig, M. (2011). Induction of human neuronal cells by defined transcription factors. *Nature*. doi: nature10202 [pii]

10.1038/nature10202 [doi]

Petters, R. M., & Sommer, J. R. (2000). Transgenic animals as models for human disease. *Transgenic Res, 9*(4-5), 347-351; discussion 345-346.

Pfeifer, A., & Hofmann, A. (2009). Lentiviral transgenesis. *Methods Mol Biol, 530*, 391-405. doi: 10.1007/978-1-59745-471-1_21

Pfeifer, A., Lim, T., & Zimmermann, K. (2010). Chapter one - Lentivirus Transgenesis. In M. W. Paul & M. S. Philippe (Eds.), *Methods in Enzymology* (Vol. Volume 477, pp. 3-15): Academic Press.

Pfisterer, U., Kirkeby, A., Torper, O., Wood, J., Nelander, J., Dufour, A., . . . Parmar, M. (2011). Direct conversion of human fibroblasts to dopaminergic neurons. *Proc Natl Acad Sci U S A, 108*(25), 10343-10348. doi: 1105135108 [pii]

10.1073/pnas.1105135108 [doi]

Phillips, T. J., Belknap, J. K., Hitzemann, R. J., Buck, K. J., Cunningham, C. L., & Crabbe, J. C. (2002). Harnessing the mouse to unravel the genetics of human disease. *Genes Brain Behav, 1*(1), 14-26.

Plomin, R., Owen, M. J., & McGuffin, P. (1994). The genetic basis of complex human behaviors. *Science, 264*(5166), 1733-1739.

Pollock, K., Stroemer, P., Patel, S., Stevanato, L., Hope, A., Miljan, E., . . . Sinden, J. D. (2006). A conditionally immortal clonal stem cell line from human cortical neuroepithelium for the treatment of ischemic stroke. *Experimental Neurology, 199*(1), 143-155. doi: http://dx.doi.org/10.1016/j.expneurol.2005.12.011

Potthoff, M. J., & Olson, E. N. (2007). MEF2: a central regulator of diverse developmental programs. *Development, 134*(23), 4131-4140. doi: 10.1242/dev.008367

Psychiatric, G. C. B. D. W. G. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet, 43*(10), 977-983. doi: 10.1038/ng.943

Pulipparacharuvil, S., Renthal, W., Hale, C. F., Taniguchi, M., Xiao, G., Kumar, A., . . . Cowan, C. W. (2008). Cocaine regulates MEF2 to control synaptic and behavioral plasticity. *Neuron, 59*(4), 621-633. doi: 10.1016/j.neuron.2008.06.020

Qu, Q., Sun, G., Murai, K., Ye, P., Li, W., Asuelime, G., . . . Shi, Y. (2013). Wnt7a Regulates Multiple Steps of Neurogenesis. *Molecular and Cellular Biology, 33*(13), 2551-2559. doi: 10.1128/mcb.00325-13

Rakic, P. (1972). Mode of cell migration to the superficial layers of fetal monkey neocortex. *J Comp Neurol, 145*(1), 61-83. doi: 10.1002/cne.901450105

Rakic, P. (1995). A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends in Neurosciences, 18*(9), 383-388. doi: http://dx.doi.org/10.1016/0166-2236(95)93934-P

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., . . . Strom, T. M. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *The Lancet, 380*(9854), 1674-1682. doi: http://dx.doi.org/10.1016/S0140-6736(12)61480-9

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., . . . Strom, T. M. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *The Lancet, 380*(9854), 1674-1682. doi: http://dx.doi.org/10.1016/S0140-6736(12)61480-9

Ravasi, T., Suzuki, H., Cannistraci, C. V., Katayama, S., Bajic, V. B., Tan, K., . . . Hayashizaki, Y. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. *Cell, 140*(5), 744-752. doi: 10.1016/j.cell.2010.01.044

Remenyi, A., Scholer, H. R., & Wilmanns, M. (2004). Combinatorial control of gene expression. *Nat Struct Mol Biol, 11*(9), 812-815. doi: 10.1038/nsmb820

Rink, E., & Wullimann, M. F. (2002). Connections of the ventral telencephalon and tyrosine hydroxylase distribution in the zebrafish brain (Danio rerio) lead to identification of an ascending dopaminergic system in a teleost. *Brain Res Bull, 57*(3-4), 385-387.

Rio, M., Royer, G., Gobin, S., de Blois, M. C., Ozilou, C., Bernheim, A., . . . Malan, V. (2013). Monozygotic twins discordant for submicroscopic chromosomal anomalies in 2p25.3 region detected by array CGH. *Clin Genet, 84*(1), 31-36. doi: 10.1111/cge.12036

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kahler, A. K., Akterin, S., . . . Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet, 45*(10), 1150-1159. doi: 10.1038/ng.2742

Robbins, P. D., & Ghivizzani, S. C. (1998). Viral Vectors for Gene Therapy. *Pharmacology & Therapeutics, 80*(1), 35-47. doi: http://dx.doi.org/10.1016/S0163-7258(98)00020-5

Romijn, H. J., Hofman, M. A., & Gramsbergen, A. (1991). At what age is the developing cerebral cortex of the rat comparable to that of the full-term newborn human baby? *Early Hum Dev, 26*(1), 61-67.

Romm, E., Nielsen, J. A., Kim, J. G., & Hudson, L. D. (2005). Myt1 family recruits histone deacetylase to regulate neural transcription. *J Neurochem, 93*(6), 1444-1453. doi: JNC3131 [pii]

10.1111/j.1471-4159.2005.03131.x [doi]

Ronan, J. L., Wu, W., & Crabtree, G. R. (2013). From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet, 14*(5), 347-359. doi: 10.1038/nrg3413

Ross, L. S., Parrett, T., & Easter, S. S., Jr. (1992). Axonogenesis and morphogenesis in the embryonic zebrafish brain. *J Neurosci, 12*(2), 467-482.

Rubenstein, J. L., & Puelles, L. (1994). Homeobox gene expression during development of the vertebrate brain. *Curr Top Dev Biol, 29*, 1-63.

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior, 2*(5), 255-267. doi: 10.1034/j.1601-183X.2003.00037.x

Salinas, P. C., & Zou, Y. (2008). Wnt signaling in neural circuit assembly. *Annu Rev Neurosci, 31*, 339-358. doi: 10.1146/annurev.neuro.31.060407.125649

Santini, S., Boore, J. L., & Meyer, A. (2003). Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. *Genome Res, 13*(6A), 1111-1122. doi: 10.1101/gr.700503

Sarnat, H. B. (2008). Disorders of segmentation of the neural tube: agenesis of selective neuromeres. *Handb Clin Neurol, 87*, 105-113. doi: 10.1016/S0072-9752(07)87007-2

Sasai, Y. (2002). Generation of dopaminergic neurons from embryonic stem cells. *Journal of Neurology, 249 Suppl 2*, Ii41-44. doi: 10.1007/s00415-002-1208-0

Sastre, M., Steiner, H., Fuchs, K., Capell, A., Multhaup, G., Condron, M. M., . . . Haass, C. (2001). Presenilin-dependent γ-secretase processing of β-amyloid precursor protein at a site corresponding to the S3 cleavage of Notch. *EMBO reports, 2*(9), 835-841. doi: 10.1093/embo-reports/kve180

Scheiffele, P., Fan, J., Choih, J., Fetter, R., & Serafini, T. (2000). Neuroligin expressed in nonneuronal cells triggers presynaptic development in contacting axons. *Cell, 101*(6), 657-669.

Scheschonka, A., Tang, Z., & Betz, H. (2007). Sumoylation in neurons: nuclear and synaptic roles? *Trends in Neurosciences, 30*(3), 85-91. doi: http://dx.doi.org/10.1016/j.tins.2007.01.003

Schmidt, R., Strahle, U., & Scholpp, S. (2013). Neurogenesis in zebrafish - from embryo to adult. *Neural Dev, 8*, 3. doi: 10.1186/1749-8104-8-3

Schurov, I. L., Handford, E. J., Brandon, N. J., & Whiting, P. J. (2004). Expression of disrupted in schizophrenia 1 (DISC1) protein in the adult and developing mouse brain indicates its role in neurodevelopment. *Mol Psychiatry, 9*(12), 1100-1110.

Sha, L., Macintyre, L., Machell, J. A., Kelly, M. P., Porteous, D. J., Brandon, N. J., . . . Pickard, B. S. Transcriptional regulation of neurodevelopmental and metabolic pathways by NPAS3. *Mol Psychiatry*. doi: mp201173 [pii]

10.1038/mp.2011.73 [doi]

Shalizi, A., Gaudilliere, B., Yuan, Z., Stegmuller, J., Shirogane, T., Ge, Q., . . . Bonni, A. (2006). A calcium-regulated MEF2 sumoylation switch controls postsynaptic differentiation. *Science, 311*(5763), 1012-1017. doi: 10.1126/science.1122513

She, H., & Mao, Z. (2011). Regulation of myocyte enhancer factor-2 transcription factors by neurotoxins. *Neurotoxicology, 32*(5), 563-566. doi: http://dx.doi.org/10.1016/j.neuro.2011.05.019

Shen, Q., Wang, Y., Dimos, J. T., Fasano, C. A., Phoenix, T. N., Lemischka, I. R., . . . Temple, S. (2006). The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat Neurosci, 9*(6), 743-751. doi: 10.1038/nn1694

Shevell, M. (2009). *Neurodevelopmental disabilities : clinical and scientific foundations*. London: Mac Keith Press for International Child Neurology Association.

Sidman, R. L., & Rakic, P. (1973). Neuronal migration, with special reference to developing human brain: a review. *Brain Res, 62*(1), 1-35.

Silberg, J., Meyer, J., Maes, H., Simonoff, E., Pickles, A., Rutter, M., . . . Eaves, L. (1996). Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37*(7), 803-816.

Skeath, J. B. (1999). At the nexus between pattern formation and cell-type specification: the generation of individual neuroblast fates in the Drosophila embryonic central nervous system. *Bioessays, 21*(11), 922-931. doi: 10.1002/(SICI)1521-1878(199911)21:11<922::AID-BIES4>3.0.CO;2-T

Smale, G., Nichols, N. R., Brady, D. R., Finch, C. E., & Horton Jr, W. E. (1995). Evidence for Apoptotic Cell Death in Alzheimer's Disease. *Experimental Neurology, 133*(2), 225-230. doi: http://dx.doi.org/10.1006/exnr.1995.1025

Smalley, S. L., Kustanovich, V., Minassian, S. L., Stone, J. L., Ogdie, M. N., McGough, J. J., . . . Nelson, S. F. (2002). Genetic Linkage of Attention-Deficit/Hyperactivity Disorder on Chromosome 16p13, in a Region Implicated in Autism. *The American Journal of Human Genetics, 71*(4), 959-963. doi: http://dx.doi.org/10.1086/342732

Smith, K. M., Daly, M., Fischer, M., Yiannoutsos, C. T., Bauer, L., Barkley, R., & Navia, B. A. (2003). Association of the dopamine beta hydroxylase gene with attention deficit hyperactivity disorder: genetic analysis of the Milwaukee longitudinal study. *Am J Med Genet B Neuropsychiatr Genet, 119b*(1), 77-85. doi: 10.1002/ajmg.b.20005

Spedding, M., Jay, T., e Silva, J. C., & Perret, L. (2005). A pathophysiological paradigm for the therapy of psychiatric disease. *Nat Rev Drug Discov, 4*(6), 467-476.

Stevanato, L., Corteling, R., Stroemer, P., Hope, A., Heward, J., Miljan, E., & Sinden, J. (2009). c-MycERTAM transgene silencing in a genetically modified human neural stem cell line implanted into MCAo rodent brain. *BMC Neuroscience, 10*(1), 86.

Stevens, S. J., van Ravenswaaij-Arts, C. M., Janssen, J. W., Klein Wassink-Ruiter, J. S., van Essen, A. J., Dijkhuizen, T., . . . Engelen, J. J. (2011). MYT1L is a candidate gene for intellectual disability in patients with 2p25.3 (2pter) deletions. *Am J Med Genet A, 155a*(11), 2739-2745. doi: 10.1002/ajmg.a.34274

Stevens, S. J. C., van Ravenswaaij-Arts, C. M. A., Janssen, J. W. H., Klein Wassink-Ruiter, J. S., van Essen, A. J., Dijkhuizen, T., . . . Engelen, J. J. M. (2011). MYT1L is a candidate gene for intellectual disability in patients with 2p25.3 (2pter) deletions. *American Journal of Medical Genetics Part A, 155*(11), 2739-2745. doi: 10.1002/ajmg.a.34274

Stevenson, J. (1992). Evidence for a genetic etiology in hyperactivity in children. *Behav Genet, 22*(3), 337-344.

Stiles, J., & Jernigan, T. L. (2010). The basics of brain development. *Neuropsychol Rev, 20*(4), 327-348. doi: 10.1007/s11065-010-9148-4

Stott, S. R. W., Metzakopian, E., Lin, W., Kaestner, K. H., Hen, R., & Ang, S.-L. (2013). Foxa1 and Foxa2 Are Required for the Maintenance of Dopaminergic Properties in Ventral Midbrain Neurons at Late Embryonic Stages. *The Journal of Neuroscience, 33*(18), 8022-8034. doi: 10.1523/jneurosci.4774-12.2013

Streisinger, G., Walker, C., Dower, N., Knauber, D., & Singer, F. (1981). Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature, 291*(5813), 293-296.

Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., . . . Hogenesch, J. B. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America, 101*(16), 6062-6067. doi: 10.1073/pnas.0400782101

Sudhof, T. C. (2008). Neuroligins and neurexins link synaptic function to cognitive disease. *Nature, 455*(7215), 903-911.

Sutcliffe, J. S. (2008). Insights into the Pathogenesis of Autism. *Science, 321*(5886), 208-209. doi: 10.1126/science.1160555

Suzuki, N., Fukushi, M., Kosaki, K., Doyle, A. D., de Vega, S., Yoshizaki, K., . . . Yamada, Y. (2012). Teneurin-4 Is a Novel Regulator of Oligodendrocyte Differentiation and Myelination of Small-Diameter Axons in the CNS. *The Journal of Neuroscience, 32*(34), 11586-11599. doi: 10.1523/jneurosci.2045-11.2012

Szatmari, P., Paterson, A. D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X. Q., . . . Meyer, K. J. (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet, 39*(3), 319-328. doi: 10.1038/ng1985

Tabuchi, K., Blundell, J., Etherton, M. R., Hammer, R. E., Liu, X., Powell, C. M., & Sudhof, T. C. (2007). A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science, 318*(5847), 71-76. doi: 10.1126/science.1146221

Takahashi, J., Palmer, T. D., & Gage, F. H. (1999). Retinoic acid and neurotrophins collaborate to regulate neurogenesis in adult-derived neural stem cell cultures. *J Neurobiol, 38*(1), 65-81.

Tamura, T., Konishi, Y., Makino, Y., & Mikoshiba, K. (1996). Mechanisms of transcriptional regulation and neural gene expression. *Neurochem Int, 29*(6), 573-581.

Tatton, W. G., Chalmers-Redman, R., Brown, D., & Tatton, N. (2003). Apoptosis in Parkinson's disease: signals for neuronal degradation. *Ann Neurol, 53 Suppl 3*, S61-70; discussion S70-62. doi: 10.1002/ana.10489

Tau, G. Z., & Peterson, B. S. (2010). Normal development of brain circuits. *Neuropsychopharmacology, 35*(1), 147-168. doi: 10.1038/npp.2009.115

Taylor, J. S., Braasch, I., Frickey, T., Meyer, A., & Van de Peer, Y. (2003). Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Research, 13*(3), 382-390.

Temple, S. (2001). The development of neural stem cells. *Nature, 414*(6859), 112-117.

Temple, S. (2001). The development of neural stem cells. *Nature, 414*(6859), 112-117. doi: 10.1038/35102174

ten Donkelaar, H. J., & Lammens, M. (2009). Development of the Human Cerebellum and Its Disorders. *Clinics in Perinatology, 36*(3), 513-530. doi: http://dx.doi.org/10.1016/j.clp.2009.06.001

Tennant-Eyles, A. J., Moffitt, H., Whitehouse, C. A., & Roberts, R. G. (2011). Characterisation of the FAM69 family of cysteine-rich endoplasmic reticulum proteins. *Biochemical and Biophysical Research Communications, 406*(3), 471-477. doi: http://dx.doi.org/10.1016/j.bbrc.2011.02.076

Thisse, B., Thisse, C. (2004). Fast Release Clones: a high throughput expression analysis. from http://zfin.org

Thisse, C., & Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc, 3*(1), 59-69. doi: nprot.2007.514 [pii]

10.1038/nprot.2007.514 [doi]

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science, 282*(5391), 1145-1147.

Tkachev, D., Mimmack, M. L., Ryan, M. M., Wayland, M., Freeman, T., Jones, P. B., . . . Bahn, S. (2003). Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. *The Lancet, 362*(9386), 798-805. doi: http://dx.doi.org/10.1016/S0140-6736(03)14289-4

Tropepe, V., & Sive, H. L. (2003). Can zebrafish be used as a model to study the neurodevelopmental causes of autism? *Genes Brain Behav, 2*(5), 268-281.

Tsai, P. T., Hull, C., Chu, Y., Greene-Colozzi, E., Sadowski, A. R., Leech, J. M., . . . Sahin, M. (2012). Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. *Nature, 488*(7413), 647-651. doi: http://www.nature.com/nature/journal/v488/n7413/abs/nature11310.html#supplementary-information

Tsygankova, O. M., Ma, C., Tang, W., Korch, C., Feldman, M. D., Lv, Y., . . . Meinkoth, J. L. (2010). Downregulation of Rap1GAP in human tumor cells alters cell/matrix and cell/cell adhesion. *Mol Cell Biol, 30*(13), 3262-3274. doi: 10.1128/mcb.01345-09

Twyman, R. (2002). Model Organisms: Fish. *The Human Genome*. Retrieved from genome.wellcome.ac.uk website:

Ullian, E. M., Sapperstein, S. K., Christopherson, K. S., & Barres, B. A. (2001). Control of synapse number by glia. *Science, 291*(5504), 657-661. doi: 10.1126/science.291.5504.657

Umemori, H., Satot, S., Yagi, T., Aizawal, S., & Yamamoto, T. (1994). Initial events of myelination involve Fyn tyrosine kinase signalling. *Nature, 367*(6463), 572-576.

Van Den Bossche, M. J., Strazisar, M., Cammaerts, S., Liekens, A. M., Vandeweyer, G., Depreeuw, V., . . . Del-Favero, J. (2013). Identification of rare copy number variants in high burden schizophrenia families. *Am J Med Genet B Neuropsychiatr Genet, 162B*(3), 273-282. doi: 10.1002/ajmg.b.32146

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol, 3*(7), Research0034.

Vascotto, S. G., Beckham, Y., & Kelly, G. M. (1997). The zebrafish's swim to fame as an experimental model in biology. *Biochem Cell Biol, 75*(5), 479-485.

Vierbuchen, T., Ostermeier, A., Pang, Z. P., Kokubu, Y., Sudhof, T. C., & Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature, 463*(7284), 1035-1041. doi: nature08797 [pii]

10.1038/nature08797 [doi]

Vierbuchen, T., & Wernig, M. (2011). Direct lineage conversions: unnatural but useful[quest]. *Nat Biotech, 29*(10), 892-907. doi: 10.1038/nbt.1946

Viti, J., Gulacsi, A., & Lillien, L. (2003). Wnt Regulation of Progenitor Maturation in the Cortex Depends on Shh or Fibroblast Growth Factor 2. *The Journal of Neuroscience, 23*(13), 5919-5927.

Volpe, J. J. (2000). Overview: normal and abnormal human brain development. *Ment Retard Dev Disabil Res Rev, 6*(1), 1-5. doi: 10.1002/(SICI)1098-2779(2000)6:1<1::AID-MRDD1>3.0.CO;2-J [pii]

10.1002/(SICI)1098-2779(2000)6:1<1::AID-MRDD1>3.0.CO;2-J

Volpe, J. J. (2008). *Neurology of the newborn* (5th ed. ed.). Philadelphia, Pa. ; London: Saunders.

Vrijenhoek, T., Buizer-Voskamp, J. E., van der Stelt, I., Strengman, E., Sabatti, C., Geurts van Kessel, A., . . . Veltman, J. A. (2008). Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am J Hum Genet, 83*(4), 504-510. doi: S0002-9297(08)00501-6 [pii]

10.1016/j.ajhg.2008.09.011 [doi]

Waites, C. L., Craig, A. M., & Garner, C. C. (2005). Mechanisms of vertebrate synaptogenesis. *Annual Review of Neuroscience, 28*, 251-274.

Walker, C., & Streisinger, G. (1983). Induction of Mutations by gamma-Rays in Pregonial Germ Cells of Zebrafish Embryos. *Genetics, 103*(1), 125-136.

Walther, W., & Stein, U. (2000). Viral vectors for gene transfer: a review of their use in the treatment of human diseases. *Drugs, 60*(2), 249-271.

Wang, S., Zhang, J., Zhao, A., Hipkens, S., Magnuson, M. A., & Gu, G. (2007). Loss of Myt1 function partially compromises endocrine islet cell differentiation and pancreatic physiological function in the mouse. *Mech Dev, 124*(11-12), 898-910. doi: S0925-4773(07)00147-5 [pii]

10.1016/j.mod.2007.08.004 [doi]

Wang, T., Zeng, Z., Li, T., Liu, J., Li, J., Li, Y., . . . Shi, Y. (2010). Common SNPs in myelin transcription factor 1-like (MYT1L): association with major depressive disorder in the Chinese Han population. *PLoS One, 5*(10), e13662. doi: 10.1371/journal.pone.0013662

Wang, Y.-X., Qian, L.-X., Yu, Z., Jiang, Q., Dong, Y.-X., Liu, X.-F., . . . Song, H.-Y. (2005). Requirements of myocyte-specific enhancer factor 2A in zebrafish cardiac contractility. *FEBS Letters, 579*(21), 4843-4850. doi: http://dx.doi.org/10.1016/j.febslet.2005.07.068

Weber, K., Bartsch, U., Stocking, C., & Fehse, B. (2008). A Multicolor Panel of Novel Lentiviral [ldquo]Gene Ontology[rdquo] (LeGO) Vectors for Functional Gene Analysis. *Mol Ther, 16*(4), 698-706.

Weber, K., Mock, U., Petrowitz, B., Bartsch, U., & Fehse, B. (2010). Lentiviral gene ontology (LeGO) vectors equipped with novel drug-selectable fluorescent proteins: new building blocks for cell marking and multi-gene analysis. *Gene Ther, 17*(4), 511-520.

Weiner, J. A., & Chun, J. (1997). Png-1, a nervous system-specific zinc finger gene, identifies regions containing postmitotic neurons during mammalian embryonic development. *J*

Comp Neurol, 381(2), 130-142. doi: 10.1002/(SICI)1096-9861(19970505)381:2<130::AID-CNE2>3.0.CO;2-4 [pii]

Wichterle, H., Lieberam, I., Porter, J. A., & Jessell, T. M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell, 110*(3), 385-397.

Wilson, S. W., & Houart, C. (2004). Early steps in the development of the forebrain. *Dev Cell, 6*(2), 167-181.

Wodarz, A., & Huttner, W. B. (2003). Asymmetric cell division during neurogenesis in Drosophila and vertebrates. *Mechanisms of Development, 120*(11), 1297-1309. doi: http://dx.doi.org/10.1016/j.mod.2003.06.003

Woods, I. G., Wilson, C., Friedlander, B., Chang, P., Reyes, D. K., Nix, R., . . . Talbot, W. S. (2005). The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Research, 15*(9), 1307-1314.

Workman, A. D., Charvet, C. J., Clancy, B., Darlington, R. B., & Finlay, B. L. (2013). Modeling Transformations of Neurodevelopmental Sequences across Mammalian Species. *The Journal of Neuroscience, 33*(17), 7368-7383. doi: 10.1523/jneurosci.5746-12.2013

Wrathall, J. R., Li, W., & Hudson, L. D. (1998). Myelin gene expression after experimental contusive spinal cord injury. *J Neurosci, 18*(21), 8780-8793.

Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., . . . Su, A. (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biology, 10*(11), R130.

Yamada, T., Yang, Y., Huang, J., Coppola, G., Geschwind, D. H., & Bonni, A. (2013). Sumoylated MEF2A coordinately eliminates orphan presynaptic sites and promotes maturation of presynaptic boutons. *J Neurosci, 33*(11), 4726-4740. doi: 10.1523/JNEUROSCI.4191-12.2013

Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., . . . Weinberg, R. A. (2004). Twist, a Master Regulator of Morphogenesis, Plays an Essential Role in Tumor Metastasis. *Cell, 117*(7), 927-939. doi: http://dx.doi.org/10.1016/j.cell.2004.06.006

Yee, K. S., & Yu, V. C. (1998). Isolation and characterization of a novel member of the neural zinc finger factor/myelin transcription factor family with transcriptional repression activity. *J Biol Chem, 273*(9), 5366-5374.

Yin, Y., She, H., Li, W., Yang, Q., Guo, S., & Mao, Z. (2012). Modulation of neuronal survival factor MEF2 by kinases in Parkinson's disease. *Frontiers in Physiology, 3*. doi: 10.3389/fphys.2012.00171

Yoo, A. S., Sun, A. X., Li, L., Shcheglovitov, A., Portmann, T., Li, Y., . . . Crabtree, G. R. (2011). MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature, 476*(7359), 228-231. doi: 10.1038/nature10323

Youn, H.-D., Sun, L., Prywes, R., & Liu, J. O. (1999). Apoptosis of T Cells Mediated by Ca2+-Induced Release of the Transcription Factor MEF2. *Science, 286*(5440), 790-793. doi: 10.1126/science.286.5440.790

Yu, Y. T., Breitbart, R. E., Smoot, L. B., Lee, Y., Mahdavi, V., & Nadal-Ginard, B. (1992). Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes Dev, 6*(9), 1783-1798.

Zeng, L., Zhang, P., Shi, L., Yamamoto, V., Lu, W., & Wang, K. (2013). Functional impacts of NRXN1 knockdown on neurodevelopment in stem cell models. *PLoS One, 8*(3), e59685. doi: 10.1371/journal.pone.0059685

Zhao, W., Zhao, S.-p., & Peng, D.-q. (2012). The effects of myocyte enhancer factor 2A gene on the proliferation, migration and phenotype of vascular smooth muscle cells. *Cell Biochemistry and Function, 30*(2), 108-113. doi: 10.1002/cbf.1823

Appendix

Table A1. All differentially expressed (FDR < 0.05) genes between *MYT1L* knock-down and non-silencing shRNA controls in SPC-04 cells. Relevant illumina probe, gene symbol and a name, log fold change (FC) for cells differentiated for 7 and 14 days, as well as uncorrected and FDR-corrected *p*-values are listed for each gene. Genes are sorted according to *p*- value; lowest to highest.

Probe ID	Gene symbol	Definition	MYT1L 7day LogFC	MYT1L 14day Log FC	p-value	FDR
520474	ODZ4	odz, odd Oz/ten-m homolog 4 (Drosophila) (ODZ4), mRNA.	-1.65	-1.97	8.27E-07	0.0039
4890181	RAP1GAP	RAP1 GTPase activating protein (RAP1GAP), mRNA.	-1.40	-2.24	2.42E-06	0.0039
5490019	GPX3	glutathione peroxidase 3 (plasma) (GPX3), mRNA.	-0.75	-1.07	1.38E-06	0.0039
6620008	KAL1	Kallmann syndrome 1 sequence (KAL1), mRNA.	-2.17	-1.94	8.19E-06	0.0088
4050025	0	mRNA; cDNA DKFZp686J23256 (from clone DKFZp686J23256)	-2.03	-1.67	7.58E-06	0.0088
3930026	RASL10A	RAS-like, family 10, member A (RASL10A), transcript variant 2, mRNA.	-1.02	-2.04	9.52E-06	0.00921
2100446	BCAN	brevican (BCAN), transcript variant 1, mRNA.	-3.19	-2.53	1.83E-05	0.0129
7050575	PPAP2B	phosphatidic acid phosphatase type 2B (PPAP2B), transcript variant 1, mRNA.	-1.52	-1.07	2.00E-05	0.0129
3370162	SPON1	spondin 1, extracellular matrix protein (SPON1), mRNA.	-0.39	-1.43	1.96E-05	0.0129
2570564	HLA-DRA	major histocompatibility complex, class II, DR alpha (HLA-DRA), mRNA.	-0.58	-1.82	2.37E-05	0.01435
6110152	C21orf62	chromosome 21 open reading frame 62 (C21orf62), mRNA.	-1.07	-1.51	2.52E-05	0.01436
110333	LOC440585	PREDICTED: hypothetical LOC440585, transcript variant 3 (LOC440585), mRNA.	-1.76	-1.67	2.93E-05	0.01522
4060433	MAOB	monoamine oxidase B (MAOB), nuclear gene encoding mitochondrial protein, mRNA.	-1.74	-2.05	3.17E-05	0.01522
3310307	C21orf63	chromosome 21 open reading frame 63 (C21orf63), mRNA.	-1.45	-1.71	3.27E-05	0.01522
3780092	TNFRSF21	tumor necrosis factor receptor superfamily, member 21 (TNFRSF21), mRNA.	-1.33	-0.70	3.30E-05	0.01522
1410403	LRRN2	leucine rich repeat neuronal 2 (LRRN2), transcript variant 2, mRNA.	-0.41	-0.92	4.35E-05	0.0191
4890021	36403	septin 3 (SEPT3), transcript variant B, mRNA.	-1.72	-1.51	5.23E-05	0.01921

6020451	RSPH1	radial spoke head 1 homolog (Chlamydomonas) (RSPH1), mRNA.	-1.28	-1.33	0.0001	0.01921
4220468	ATP1B1	ATPase, Na+/K+ transporting, beta 1 polypeptide (ATP1B1), transcript variant 2, mRNA.	-1.08	-1.01	9.53E-05	0.01921
3520039	NLGN3	neuroligin 3 (NLGN3), mRNA.	-1.07	-0.85	8.87E-05	0.01921
1980246	MYO5C	myosin VC (MYO5C), mRNA.	-0.98	-0.90	9.68E-05	0.01921
3850059	LMO2	LIM domain only 2 (rhombotin-like 1) (LMO2), mRNA.	-0.97	-1.65	7.73E-05	0.01921
2600465	SH3GL2	SH3-domain GRB2-like 2 (SH3GL2), mRNA.	-0.97	-1.08	5.38E-05	0.01921
6620538	UBL3	ubiquitin-like 3 (UBL3), mRNA.	-0.96	-1.07	6.04E-05	0.01921
2760239	RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated) (RASGRP1), mRNA.	-0.95	-1.10	6.05E-05	0.01921
6900630	ATP2B4	ATPase, Ca++ transporting, plasma membrane 4 (ATP2B4), transcript variant 2, mRNA.	-0.94	-0.70	7.39E-05	0.01921
1230521	0	mRNA; cDNA DKFZp686N1644 (from clone DKFZp686N1644)	-0.90	-1.01	7.48E-05	0.01921
6100482	ATP2B4	ATPase, Ca++ transporting, plasma membrane 4 (ATP2B4), transcript variant 1, mRNA.	-0.83	-0.79	9.12E-05	0.01921
4210411	NDRG2	NDRG family member 2 (NDRG2), transcript variant 6, mRNA.	-0.73	-1.00	9.00E-05	0.01921
6040500	WDR16	WD repeat domain 16 (WDR16), transcript variant 2, mRNA.	-0.72	-1.16	0.0001	0.01921
7050333	HEPACAM	hepatocyte cell adhesion molecule (HEPACAM), mRNA.	-0.68	-1.56	5.28E-05	0.01921
6060482	ATP1B1	ATPase, Na+/K+ transporting, beta 1 polypeptide (ATP1B1), transcript variant 1, mRNA.	-0.66	-0.92	9.03E-05	0.01921
7320041	CHST15	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15 (CHST15), mRNA.	-0.60	-1.07	6.32E-05	0.01921
110239	CALCR	calcitonin receptor (CALCR), mRNA.	-0.56	-0.98	8.81E-05	0.01921
1690403	PMP2	peripheral myelin protein 2 (PMP2), mRNA.	-0.36	-1.40	5.93E-05	0.01921
6040451	SLC47A2	solute carrier family 47, member 2 (SLC47A2), transcript variant 1, mRNA.	-0.15	-1.15	6.29E-05	0.01921
4780563	GABBR2	gamma-aminobutyric acid (GABA) B receptor, 2 (GABBR2), mRNA.	-1.81	-0.73	0.00011	0.01938
5960475	RHPN2	rhophilin, Rho GTPase binding protein 2 (RHPN2), mRNA.	-1.24	-1.34	0.00011	0.01938
7320072	DTNA	dystrobrevin, alpha (DTNA), transcript variant 8, mRNA.	-0.82	-1.11	0.00012	0.02077

2480338	DBI	diazepam binding inhibitor (GABA receptor modulator, acyl-Coenzyme A binding protein) (DBI), mRNA.	-0.43	-0.99	0.00013	0.02175
4760747	TPST1	tyrosylprotein sulfotransferase 1 (TPST1), mRNA.	-0.90	-1.16	0.00014	0.02179
5360689	PROX1	prospero homeobox 1 (PROX1), mRNA.	-0.70	-0.68	0.00014	0.02181
5360301	LMO3	LIM domain only 3 (rhombotin-like 2) (LMO3), transcript variant 1, mRNA.	-1.53	-1.15	0.00015	0.02261
4780044	LOC389386	PREDICTED: misc_RNA (LOC389386), partial miscRNA.	-0.52	-0.99	0.00015	0.02261
870048	PHYHIPL	phytanoyl-CoA 2-hydroxylase interacting protein-like (PHYHIPL), mRNA.	-1.77	-0.95	0.00017	0.02434
940630	MOBKL2B	MOB1, Mps One Binder kinase activator-like 2B (yeast) (MOBKL2B), mRNA.	-1.26	-0.90	0.00018	0.02434
4180452	ATP1B1	ATPase, Na+/K+ transporting, beta 1 polypeptide (ATP1B1), transcript variant 1, mRNA.	-0.90	-1.05	0.00017	0.02434
460767	BGN	biglycan (BGN), mRNA.	-0.56	-1.02	0.00018	0.02434
6100356	ALPL	alkaline phosphatase, liver/bone/kidney (ALPL), transcript variant 1, mRNA.	-0.24	-0.91	0.00017	0.02434
6550070	LPAR4	lysophosphatidic acid receptor 4 (LPAR4), mRNA.	-0.68	-0.99	0.00018	0.02446
540326	NCAN	neurocan (NCAN), mRNA.	-1.90	-1.01	0.0002	0.02578
2320047	NFIA	nuclear factor I/A (NFIA), mRNA.	-1.79	-1.41	0.00024	0.02578
3190608	SLC6A9	solute carrier family 6 (neurotransmitter transporter, glycine), member 9 (SLC6A9), transcript variant 3, mRNA.	-1.29	-1.46	0.00025	0.02578
2600554	BTBD17	BTB (POZ) domain containing 17 (BTBD17), mRNA.	-1.08	-1.44	0.00023	0.02578
4180047	C1orf194	chromosome 1 open reading frame 194 (C1orf194), mRNA.	-0.99	-1.30	0.00024	0.02578
7550358	NELL2	NEL-like 2 (chicken) (NELL2), mRNA.	-0.97	-1.66	0.00023	0.02578
4570253	DTNA	dystrobrevin, alpha (DTNA), transcript variant 6, mRNA.	-0.90	-1.15	0.00025	0.02578
4560707	FAM183A	family with sequence similarity 183, member A (FAM183A), mRNA.	-0.86	-1.04	0.00025	0.02578
4210678	PCSK1N	proprotein convertase subtilisin/kexin type 1 inhibitor (PCSK1N), mRNA.	-0.76	-1.12	0.00022	0.02578
1090326	TIMP4	TIMP metallopeptidase inhibitor 4 (TIMP4), mRNA.	-0.76	-0.99	0.00021	0.02578
1710020	SMOC1	SPARC related modular calcium binding 1 (SMOC1), transcript variant 1, mRNA.	-0.72	-1.32	0.00021	0.02578

1570487	WDR16	WD repeat domain 16 (WDR16), transcript variant 2, mRNA.	-0.66	-1.05	0.00021	0.02578
1980484	EFHD1	EF-hand domain family, member D1 (EFHD1), mRNA.	-0.63	-1.26	0.0002	0.02578
4230195	NAV2	neuron navigator 2 (NAV2), transcript variant 2, mRNA.	-0.62	-1.26	0.00022	0.02578
2900626	SRI	sorcin (SRI), transcript variant 2, mRNA.	-0.62	-1.15	0.00025	0.02578
270168	HLA-DRA	major histocompatibility complex, class II, DR alpha (HLA-DRA), mRNA.	-0.27	-1.28	0.00024	0.02578
5860039	CALB2	calbindin 2 (CALB2), transcript variant CALB2c, mRNA.	-1.19	-0.62	0.00025	0.02586
1660019	SNCAIP	synuclein, alpha interacting protein (SNCAIP), mRNA.	-1.54	-1.25	0.00027	0.02606
6220097	PPAP2B	phosphatidic acid phosphatase type 2B (PPAP2B), transcript variant 2, mRNA.	-1.02	-0.68	0.00026	0.02606
4780463	RAB6B	RAB6B, member RAS oncogene family (RAB6B), mRNA.	-0.95	-1.29	0.00027	0.02606
2450367	LRIG1	leucine-rich repeats and immunoglobulin-like domains 1 (LRIG1), mRNA.	-0.86	-1.07	0.00027	0.02606
6770392	SRI	sorcin (SRI), transcript variant 1, mRNA.	-0.68	-1.20	0.00026	0.02606
240086	PHGDH	phosphoglycerate dehydrogenase (PHGDH), mRNA.	-0.54	-1.30	0.00026	0.02606
6040066	36403	septin 3 (SEPT3), transcript variant B, mRNA.	-1.50	-1.21	0.00029	0.02706
2750176	AQP4	aquaporin 4 (AQP4), transcript variant a, mRNA.	-1.57	-1.97	0.00032	0.02909
130754	PTPRZ1	protein tyrosine phosphatase, receptor-type, Z polypeptide 1 (PTPRZ1), mRNA.	-0.85	-0.99	0.00032	0.02909
6550681	IGSF5	immunoglobulin superfamily, member 5 (IGSF5), mRNA.	-0.80	-0.85	0.00032	0.02909
150474	CA12	carbonic anhydrase XII (CA12), transcript variant 1, mRNA.	-0.64	-1.88	0.00031	0.02909
6220397	SC5DL	sterol-C5-desaturase (ERG3 delta-5-desaturase homolog, S. cerevisiae)-like (SC5DL), transcript variant 1, mRNA.	-0.51	-0.87	0.00032	0.02936
1110048	SCRG1	scrapie responsive protein 1 (SCRG1), mRNA.	-1.08	-0.31	0.00033	0.02936
7000685	C1orf192	chromosome 1 open reading frame 192 (C1orf192), mRNA.	-1.15	-0.83	0.00034	0.02965
520324	REEP5	receptor accessory protein 5 (REEP5), mRNA.	-0.54	-0.72	0.00035	0.0297
4610672	FLJ14213	protor-2 (FLJ14213), mRNA.	-0.57	-0.92	0.00035	0.02991
360296	CRMP1	collapsin response mediator protein 1 (CRMP1), transcript variant 1, mRNA.	-1.45	-0.66	0.00037	0.03074

5050093	KBTBD11	kelch repeat and BTB (POZ) domain containing 11 (KBTBD11), mRNA.	-0.87	-1.16	0.00038	0.03189
4210746	PACRG	PARK2 co-regulated (PACRG), transcript variant 3, mRNA.	-1.00	-0.98	0.00039	0.03191
2940189	LRRC4C	leucine rich repeat containing 4C (LRRC4C), mRNA.	-0.57	-0.57	0.0004	0.03234
2120707	CETN2	centrin, EF-hand protein, 2 (CETN2), mRNA.	-0.81	-0.76	0.0004	0.03254
1850242	ARMC3	armadillo repeat containing 3 (ARMC3), mRNA.	-0.70	-1.03	0.00043	0.03354
5490768	GPR56	G protein-coupled receptor 56 (GPR56), transcript variant 3, mRNA.	-1.22	-0.98	0.00043	0.0337
1990079	FBXO32	F-box protein 32 (FBXO32), transcript variant 2, mRNA.	-1.57	-0.42	0.00045	0.03461
4590102	WRB	tryptophan rich basic protein (WRB), mRNA.	-0.73	-0.47	0.00045	0.03461
3940095	NAV2	neuron navigator 2 (NAV2), transcript variant 2, mRNA.	-0.27	-0.83	0.00047	0.0358
6420630	SCARA3	scavenger receptor class A, member 3 (SCARA3), transcript variant 1, mRNA.	-0.32	-0.85	0.00049	0.03706
7330544	ALDOC	aldolase C, fructose-bisphosphate (ALDOC), mRNA.	-0.42	-1.75	0.0005	0.03723
4220672	MT1F	metallothionein 1F (MT1F), mRNA.	-1.62	-1.08	0.00051	0.03752
1710736	DOCK10	dedicator of cytokinesis 10 (DOCK10), mRNA.	-1.06	-0.70	0.00053	0.03849
6020682	RGMA	RGM domain family, member A (RGMA), mRNA.	-2.01	-1.78	0.00053	0.0387
4290097	CD99	CD99 molecule (CD99), transcript variant 1, mRNA.	-0.59	-0.90	0.00056	0.03986
2630519	RSHL3	radial spokehead-like 3 (RSHL3), mRNA.	-1.59	-1.46	0.00058	0.04023
2850301	AGT	angiotensinogen (serpin peptidase inhibitor, clade A, member 8) (AGT), mRNA.	-1.28	-1.26	0.00058	0.04023
1090246	BCAN	brevican (BCAN), transcript variant 2, mRNA.	-1.03	-0.68	0.00059	0.04023
5960180	PRPH	peripherin (PRPH), mRNA.	-0.92	-1.49	0.00059	0.04023
1190673	LMAN2L	lectin, mannose-binding 2-like (LMAN2L), mRNA.	-0.91	-0.42	0.0006	0.04034
4490180	BEXL1	PREDICTED: brain expressed X-linked-like 1 (BEXL1), mRNA.	-0.70	-0.97	0.0006	0.04053
270022	CEP135	centrosomal protein 135kDa (CEP135), mRNA.	-0.94	-0.93	0.00061	0.0406
1570184	KCNN3	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3 (KCNN3), transcript variant 1, mRNA.	-0.83	-0.90	0.00062	0.04085
6650070	LIX1	Lix1 homolog (mouse) (LIX1), mRNA.	-1.23	-1.03	0.00067	0.04352

2900348	ZFP36	zinc finger protein 36, C3H type, homolog (mouse) (ZFP36), mRNA.	-1.08	-1.47	0.00067	0.04352
4040471	RUFY3	RUN and FYVE domain containing 3 (RUFY3), transcript variant 2, mRNA.	-1.05	-0.72	0.00067	0.04352
1990333	ASTN1	astrotactin 1 (ASTN1), transcript variant 1, mRNA.	-0.33	-0.83	0.00066	0.04352
4780187	0	cDNA clone IMAGE:5263177	-0.84	-0.74	0.00068	0.04353
2900390	VCAM1	vascular cell adhesion molecule 1 (VCAM1), transcript variant 1, mRNA.	-0.80	-1.24	0.0007	0.04384
6420168	DBNDD2	dysbindin (dystrobrevin binding protein 1) domain containing 2 (DBNDD2), transcript variant 3, mRNA.	-0.78	-0.99	0.00071	0.04395
2070161	AQP4	aquaporin 4 (AQP4), transcript variant a, mRNA.	-0.49	-0.81	0.00073	0.04447
270201	MEGF10	multiple EGF-like-domains 10 (MEGF10), mRNA.	-0.97	-0.92	0.00074	0.04477
870632	0	clone 23700 mRNA sequence	-0.80	-0.79	0.00075	0.0451
7320139	SOX8	SRY (sex determining region Y)-box 8 (SOX8), mRNA.	-1.01	-1.13	0.0008	0.04697
290239	NAV2	neuron navigator 2 (NAV2), transcript variant 2, mRNA.	-0.55	-1.43	0.00079	0.04697
5360370	SNTG1	syntrophin, gamma 1 (SNTG1), mRNA.	-0.55	-0.69	0.00082	0.04771
3610735	F12	coagulation factor XII (Hageman factor) (F12), mRNA.	-0.43	-0.64	0.00086	0.04952
4890241	GPR56	G protein-coupled receptor 56 (GPR56), transcript variant 2, mRNA.	-1.24	-1.30	0.00088	0.04994
2810026	WNT7A	wingless-type MMTV integration site family, member 7A (WNT7A), mRNA.	-0.66	-0.45	0.00087	0.04994
1440014	MYO5B	myosin VB (MYO5B), mRNA.	-0.59	-0.70	0.00089	0.05022
4220068	CD70	CD70 molecule (CD70), mRNA.	1.36	1.11	2.33E-06	0.0039
5270367	CTSC	cathepsin C (CTSC), transcript variant 1, mRNA.	1.48	0.49	2.01E-06	0.0039
3450138	CTSC	cathepsin C (CTSC), transcript variant 1, mRNA.	1.66	0.80	8.90E-07	0.0039
4290403	CMTM7	CKLF-like MARVEL transmembrane domain containing 7 (CMTM7), transcript variant 2, mRNA.	0.86	0.80	4.63E-06	0.00639
6250019	SHC1	SHC (Src homology 2 domain containing) transforming protein 1 (SHC1), transcript variant 2, mRNA.	1.11	1.25	1.35E-05	0.01186
670386	ID1	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein (ID1), transcript variant 2, mRNA.	1.27	1.67	1.82E-05	0.0129

7000577	GYPC	glycophorin C (Gerbich blood group) (GYPC), transcript variant 2, mRNA.	0.78	1.39	5.60E-05	0.01921
2260196	LOC392437	PREDICTED: misc_RNA (LOC392437), miscRNA.	0.80	0.69	8.94E-05	0.01921
7320551	LYN	v-yes-1 Yamaguchi sarcoma viral related oncogene homolog (LYN), mRNA.	0.82	0.49	7.00E-05	0.01921
3830092	RGS10	regulator of G-protein signaling 10 (RGS10), transcript variant 1, mRNA.	0.93	0.73	0.00011	0.01921
990300	CALCRL	calcitonin receptor-like (CALCRL), mRNA.	0.98	1.00	0.0001	0.01921
7160253	LAMA4	laminin, alpha 4 (LAMA4), mRNA.	1.08	0.77	7.55E-05	0.01921
6560301	SH2B3	SH2B adaptor protein 3 (SH2B3), mRNA.	1.17	1.78	9.13E-05	0.01921
2640544	ASS1	argininosuccinate synthetase 1 (ASS1), transcript variant 2, mRNA.	2.14	2.21	0.0001	0.01921
110433	ASS1	argininosuccinate synthetase 1 (ASS1), transcript variant 1, mRNA.	2.30	2.36	6.70E-05	0.01921
610152	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila) (DAB2), mRNA.	0.66	0.73	0.00011	0.01938
2710730	PLS3	plastin 3 (T isoform) (PLS3), mRNA.	1.44	1.29	0.00012	0.01991
6020286	FST	follistatin (FST), transcript variant FST344, mRNA.	0.39	1.47	0.00014	0.02258
6250192	COL6A3	collagen, type VI, alpha 3 (COL6A3), transcript variant 3, mRNA.	0.28	1.02	0.00018	0.02434
6020152	SH3KBP1	SH3-domain kinase binding protein 1 (SH3KBP1), transcript variant 1, mRNA.	0.60	0.93	0.00018	0.02446
2120053	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1 (CYP1B1), mRNA.	0.60	1.73	0.00022	0.02578
7210768	0	17000531886861 GRN_ES cDNA 5, mRNA sequence	0.76	0.39	0.00024	0.02578
2340241	IMPA2	inositol(myo)-1(or 4)-monophosphatase 2 (IMPA2), mRNA.	0.91	0.63	0.00022	0.02578
2060040	ADORA2B	adenosine A2b receptor (ADORA2B), mRNA.	0.97	0.73	0.0002	0.02578
2190603	PLS3	plastin 3 (T isoform) (PLS3), mRNA.	1.38	1.12	0.0002	0.02578
6040097	ARAP3	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 3 (ARAP3), mRNA.	0.77	0.67	0.00027	0.02606
6580270	LOC646723	PREDICTED: similar to Keratin, type I cytoskeletal 18 (Cytokeratin-18) (CK-18) (Keratin-18) (K18) (LOC646723), mRNA.	1.39	0.61	0.00033	0.02936

7210398	GYPC	glycophorin C (Gerbich blood group) (GYPC), transcript variant 2, mRNA.	0.73	1.31	0.00035	0.0297
5130435	LAMA4	laminin, alpha 4 (LAMA4), mRNA.	1.05	1.03	0.00034	0.0297
3930367	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila) (DAB2), mRNA.	0.65	0.85	0.00039	0.03191
2640768	CTSC	cathepsin C (CTSC), transcript variant 1, mRNA.	0.72	0.14	0.00041	0.0326
7610433	SLC1A5	solute carrier family 1 (neutral amino acid transporter), member 5 (SLC1A5), mRNA.	0.79	1.13	0.00041	0.0326
5050053	TXNDC5	thioredoxin domain containing 5 (endoplasmic reticulum) (TXNDC5), transcript variant 1, mRNA.	0.61	0.69	0.0005	0.03738
3450059	SH3KBP1	SH3-domain kinase binding protein 1 (SH3KBP1), transcript variant 1, mRNA.	0.49	0.82	0.00055	0.03926
5720136	RAB34	RAB34, member RAS oncogene family (RAB34), mRNA.	1.14	0.93	0.00055	0.03926
7160343	MYLK	myosin light chain kinase (MYLK), transcript variant 8, mRNA.	0.58	2.34	0.00058	0.04023
6520064	TMEM178	transmembrane protein 178 (TMEM178), mRNA.	1.38	1.73	0.00059	0.04023
4070356	FLG	filaggrin (FLG), mRNA.	1.74	0.92	0.00069	0.04384
4570255	LEF1	lymphoid enhancer-binding factor 1 (LEF1), mRNA.	0.85	0.91	0.0007	0.04384
2630195	VAMP5	vesicle-associated membrane protein 5 (myobrevin) (VAMP5), mRNA.	1.41	1.44	0.0007	0.04384
3440338	RNF181	ring finger protein 181 (RNF181), mRNA.	0.75	0.87	0.00072	0.04395
5360273	SNCA	synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant NACP112, mRNA.	0.76	0.67	0.00071	0.04395
1070121	LOC340274	PREDICTED: misc_RNA (LOC340274), miscRNA.	1.49	1.68	0.00078	0.04637
2690538	LOC341230	PREDICTED: misc_RNA (LOC341230), miscRNA.	1.02	1.01	0.00081	0.04754
5130438	PIPOX	pipecolic acid oxidase (PIPOX), mRNA.	1.40	0.68	0.00083	0.04779